"Express Mail" mailing label number EV 314842865 US

Date of Deposit 1 October 2003.

22386 U.S. PTO 10/677669 100103

REQUEST FOR FILING A CONTINUING PATENT APPLICATION UNDER 37 CFR § 1.53(b)(1)

Case No.	ANTICIPATED CLASSIFICATION OF THIS APPLICATION		PRIOR APPLICATION EXAMINER	ART UNIT
	CLASS	SUBCLASS		
10466/485			Elizabeth Kemmerer	1646

Address to:
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

This is a request for filing a ☒ continuation ☐ divisional application under 37 CFR § 1.53(b)(1), of pending prior application number 09/944,396, filed on August 30, 2001, entitled SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME.

1. ☒ Copy Of the Prior application, including 34 sheets of drawings, 152 pages of Application

2. ☒ Copy of the Declaration filed in the Prior application.

3. ☐ PTO Form 1449 and Information Disclosure Statement.

CLAIMS	(1) FOR	(2) NUMBER FILED	(3) NUMBER EXTRA	(4) RATE	(5) CALCULATIONS
	TOTAL CLAIMS (37 CFR 1.16(c))	20 - 20 =	0	x $ 18 =	$
	INDEPENDENT CLAIMS (37 CFR 1.16(b))	3 - 3 =	0	x $ 84 =	$
	MULTIPLE DEPENDENT CLAIMS (if applicable) (37 CFR 1.16(d))			+ $280 =	$ 0
			BASIC FEE (37 CFR 1.16(a))		$ 770
			Total of above Calculations =		$
	Reduction by 50% for filing by small entity (Note 37 CFR 1.9, 1.27, 1.28)				$
				TOTAL =	$770.00

4. ☐ A verified statement to establish small entity status under 37 CFR 1.9 and 1.27

   ☐ is enclosed.

   ☐ was filed in prior application number ____ and such status is still proper and desired (37 CFR 1.28(a)).

5. ☒ The Commissioner is hereby authorized to charge any fees which may be required under 37 CFR 1.16 and 1.17, or credit any overpayment to Deposit Account No. 23-1925. A duplicate copy of this sheet is enclosed.

6. ☒ Enclosed is a check for $ 770.00 to cover the filing fees.

7. ☒ Cancel in this application original claims 1-21 of the prior application and otherwise enter the attached preliminary amendment before calculating the filing fee.

8. ☒ The inventor(s) of the invention being claimed in this application is(are): Kevin P. Baker, David Botstein, Dan L. Eaton, Napoleone Ferrara, Ellen Filvaroff, Mary E. Gerritsen, Audrey Goddard, Paul J. Godowski, J. Christopher Grimaldi, Austin I. Gurney, Kenneth J. Hillan, Ivar J. Kljavin, Mary A. Napier, Margaret Ann Roy, Daniel Tumas, William I. Wood.

9. ☐ This application is being filed by less than all the inventors named in the prior application. In accordance with 37 CFR 1.63(d)(2), the Commissioner is requested to delete the name(s) of the following person or persons who are not inventors of the invention being claimed in this application: ____.

10. ☒ Amend the specification by inserting before the first line the sentence: "This application is a ☒ continuation ☐ division of application number 09/944,396, filed This application is a continuation of, and claims priority under 35 USC Section 120 to, US Application 09/866,028 filed 5/25/2001, which is a continuation of, and claims priority under 35 USC Section 120 to, PCT Application PCT/US99/28301 filed 12/1/1999, which claims priority under 35 USC Section 119 to US Provisional Application 60/113,296 filed 12/22/1998, where PCT/US99/28301 is a continuation-in-part of, and claims priority under 35 USC Section 120 to, US Application 09/254,311 filed 3/3/1999, now abandoned, which is the National Stage filed under 35 USC Section 371 of PCT Application PCT/US98/25108 filed 12/1/1998, which claims priority under 35 USC Section 119 to US Provisional Application 60/069,873 filed 12/17/1997."

11. ☐ New formal drawings are enclosed.

12. ☐ Priority of foreign application number _____, filed on _____ respectively in _____ is claimed under 35 U.S.C. 119.

    ☐ The certified copy has been filed in prior application number _____, filed

13. ☒ A preliminary amendment is enclosed.

14. ☒ The prior application is assigned of record to GENENTECH, INC.

15. ☒ Also enclosed: Request To Use Computer Readable Form Of Sequence Listing From Parent Application Pursuant to 37 C.F.R. Section 1.821(e); and Copy of Sequence Listing (72) pages.

16. ☒ The power of attorney in the prior application is to: K. Shannon Mrksich, Reg. No. 36,675 and other attorneys at the firm of BRINKS HOFER GILSON & LIONE.

    a. ☒ The power appears in the original papers in the prior application.

    b. ☐ Since the power does not appear in the original papers, a copy of the power in the prior application is enclosed.

    c. ☒ Address all future correspondence to: (may only be completed by applicant, or attorney or agent of record.)

K. Shannon Mrksich
BRINKS HOFER GILSON & LIONE
P.O. BOX 10395
CHICAGO, IL 60610
(312)321-4200

10-1-03
Date

S. Mrksich
Signature
Name: K. Shannon Mrksich
Reg. No. 36,675

☐ Inventor(s)
☐ Assignee of complete interest
☒ Attorney or agent of record
☐ Filed under 37 CFR 1.34(a)
Registration Number if acting under 37 CFR 1.34(a): _____.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPL NO.	FILING OR 371 (c) DATE	ART UNIT	FIL FEE REC'D	ATTY.DOCKET NO	DRAWINGS	TOT CLMS	IND CLMS
10/677,669	10/01/2003	1646	770	10466/485	34	5	1

K. Shannon Mrksich
BRINKS HOFER GILSON & LIONE
P.O. BOX 10395
CHICAGO, IL 60610

CONFIRMATION NO. 1088

FILING RECEIPT

OC000000011429949

Date Mailed: 12/05/2003

Receipt is acknowledged of this regular Patent Application. It will be considered in its order and you will be notified as to the results of the examination. Be sure to provide the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION when inquiring about this application. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please write to the Office of Initial Patent Examination's Filing Receipt Corrections, facsimile number 703-746-9195. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections (if appropriate).**

Applicant(s)

Kevin P. Baker, Darnestown, MD;
David Botstein, Belmont, CA;
Dan L. Eaton, San Rafael, CA;
Napoleone Ferrara, San Francisco, CA;
Ellen Filvaroff, San Francisco, CA;
Mary E. Gerritsen, San Mateo, CA;
Audrey Goddard, San Francisco, CA;
Paul J. Godowski, Burlingame, CA;
J. Christopher Grimaldi, San Francisco, CA;
Austin L. Gurney, Belmont, CA;
Kenneth J. Hillan, San Francisco, CA;
Ivar J. Kljavin, Pacifica, CA;
Mary A. Napier, Hillsborough, CA;
Margaret Ann Roy, San Francisco, CA;
Daniel Tumas, Orinda, CA;
William I. Wood, Hillsborough, CA;

Assignment For Published Patent Application

GENENTECH, INC.;

Domestic Priority data as claimed by applicant

This application is a CON of 09/944,396 08/30/2001
and is a CON of 09/866,028 05/25/2001 PAT 6,642,360
which is a CON of PCT/US99/28301 12/01/1999
which claims benefit of 60/113,296 12/22/1998
and is a CIP of 09/254,311 03/03/1999 ABN

which is a 371 of PCT/US98/25108 12/01/1998
which claims benefit of 60/069,873 12/17/1997

Foreign Applications

If Required, Foreign Filing License Granted: 12/05/2003

Projected Publication Date: 03/18/2004

Non-Publication Request: No

Early Publication Request: No

Title

Secreted and transmembrane polypeptides and nucleic acids encoding the same

Preliminary Class

530

LICENSE FOR FOREIGN FILING UNDER
Title 35, United States Code, Section 184
Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Office of Export Administration, Department of Commerce (15 CFR 370.10 (j)); the Office of Foreign Assets Control, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy

order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

United States Patent and Trademark Office OG Notices: 18 March 2003

Claiming the Benefit of a Prior-Filed Application under 35 U.S.C. 119(e), 120, 121, and 365(c)

Summary: This notice clarifies how benefit claims under 35 U.S.C. 119(e), 120, 121 and 365(c) must be presented in applications in order to be in compliance with the relevant statute and patent regulations, and accepted by the United States Patent and Trademark Office (Office).

35 U.S.C. 120 Benefit Claims

Benefit claims under 35 U.S.C. 120 must include a specific reference to the earlier filed (nonprovisional) application for which a benefit is sought. A "specific reference" requires: (1) the identification of the prior (nonprovisional) application by application number; and (2) an indication of the relationship between the nonprovisional applications, except for the benefit claim to the prior application in a continued prosecution application (CPA). The relationship between any two nonprovisional applications will be an indication that the later-filed nonprovisional application is either a continuation, divisional, or continuation-in-part of the prior-filed nonprovisional application. When there are benefit claims to multiple prior nonprovisional applications (e.g., a string of prior nonprovisional applications), the relationship must include an identification of each nonprovisional application as either a continuation, divisional, or continuation-in-part application of a specific prior nonprovisional application for which a benefit is claimed. The identification is needed in order to be able to verify if copendency exists throughout the entire chain of prior nonprovisional applications.

35 U.S.C. 119(e) Benefit Claims

Benefit claims under 35 U.S.C. 119(e) must include a specific reference to the earlier filed provisional application for which a benefit is sought. A specific reference in this situation, however, only requires identification of the prior provisional application by the application number. No relationship between the subject nonprovisional application and the prior provisional application should be specified. If the subject nonprovisional application was not filed within twelve months of the filing date of the provisional application, the subject application must also include a benefit claim under 35 U.S.C. 120 to an intermediate prior nonprovisional application that was filed within twelve months of the filing date of the provisional application. Therefore, in addition to the identification of the provisional application, the proper benefit claim for this situation must also identify the intermediate prior nonprovisional application that is directly claiming the benefit of the provisional application, and indicate the relationship between the nonprovisional applications (e.g., an indication that the subject application is a continuation of the intermediate prior nonprovisional application).

Statement of the Problem

Background: Some applicants have been submitting patent applications which include, or are amended to include, at the beginning of the specification, a statement that benefits are claimed under 35 U.S.C. 119(e) and/or 120 to prior applications, followed by a listing of many prior nonprovisional and provisional applications. The listings do not indicate: (1) the specific relationship (i.e., continuation, divisional, or continuation-in-part) between the nonprovisional applications, as required by 37 CFR 1.78(a)(2)(i); and/or (2) each nonprovisional application which is directly claiming priority to a provisional application. Without such information, the Office does not have sufficient information to enter the benefit claims into the Office's computer database.

When entering benefit claims for an application into the Office's database, the relationship (i.e., continuation, divisional or continuation-in-part) between the nonprovisional applications is required. Further, the system will not accept any benefit claim to a provisional application if the provisional application was filed more than twelve months prior to the filing date of the subject application, unless the applicant clearly identifies, and claims the benefit of, a prior nonprovisional application that was filed within twelve months of the filing date of the provisional application. Accordingly, if benefit claims are presented without all the required information, the Office will not be able to enter such benefit claims into the Office's database, the filing receipt will not reflect the prior application(s), and the projected publication date will not be calculated as a function of an earlier application's filing date.

The specific relationships between (each of) the nonprovisional applications in a chain of nonprovisional applications are also important because such information will be printed in the application publication, and/or patent. Furthermore, the designation of an application as a continuation (rather than as a continuation-in-part) is an indication that the entire invention claimed in an application has support in the prior application, whereas the designation of an application as a continuation-in-part is an indication that the claimed invention is not entirely supported by the prior application. Thus, the specific relationship between nonprovisional applications in a chain of benefit claims, and the indication of the specific nonprovisional application(s) that is directly claiming the benefit of a provisional application, will provide the information that is needed by examiners and the public in order to determine the effective prior art date of the application publication, or patent, under 35 U.S.C. 102(e).

When benefit claims are required to, but do not, include: (1) an identification of (all) intermediate benefit claims, and/or (2) the relationship between nonprovisional applications, the Office may not be able to publish applications promptly after the expiration of a period of eighteen months from the earliest filing date for which a benefit is sought under title 35, United States Code (eighteen-month publication), nor have the accuracy desired of such benefit claims in application publications. Further, the objection (by the Office), correction (by applicant), and review/entry of changes (by the Office) cycle for non-compliant benefit claims is a burdensome effort on both applicants and the Office that can be totally avoided if such benefit claims are properly submitted the first time. Accordingly, it is hoped that applicants will submit benefit claims with all the required information as set forth in this notice and, correspondingly, avoid submitting non-compliant benefit claims that leads to extra work for both the Office and applicants.

Procedures for Making Proper Benefit Claims

Part I: Reference to Prior Nonprovisional Application(s) Per 37 CFR 1.78(a)(2)(i) Must Include Identification of, and Relationship Between, Applications

35 U.S.C. 120 provides that no application shall be entitled to the benefit of an earlier filed application unless it contains, or is amended to contain, a specific reference to the earlier filed application. The specific reference required by 35 U.S.C. 120 is the reference required by 37 CFR 1.78(a)(2). 37 CFR 1.78(a)(2)(i) requires that any nonprovisional application that claims the benefit of one or more prior-filed copending nonprovisional applications, or international applications designating the United States, must contain, or be amended to contain, a reference to each such prior-filed application, identifying it by application number (consisting of the series code and serial number), or international application number and international filing date, and indicating the relationship of the applications. 37 CFR 1.78(a)(2)(iv) also provides that a request for a continued prosecution application (CPA) under 37 CFR 1.53(d) is the specific reference required under 35 U.S.C. 120 to the prior-filed application. Therefore, except for the benefit claim to the prior-filed application in a CPA, benefit claims under 35 U.S.C. 120, including claims under 35 U.S.C. 121 and 365(c), must not only identify the earlier application by application number, or by international application number and international filing date, but they must also indicate the relationship between the applications.

Examples

The relationship between the applications is whether the subject application is a continuation, divisional, or continuation-in-part of a prior-filed nonprovisional application. An example of a proper benefit claim is: "This application is a continuation of Application No. 10/ - , filed - ." A benefit claim that merely states: "This application claims the benefit of Application No. 10/ - , filed - ." does not comply with 37 CFR 1.78(a)(2)(i), since the relationship between the applications is not stated. In addition, a benefit claim that merely states: "This application is a continuing application of Application No. 10/ - , filed - ." does not comply with 37 CFR 1.78(a)(2)(i) since the proper relationship, which includes the type of continuing application, is not stated. It is also noted that the status of each nonprovisional parent application (if it is patented or abandoned) should also be indicated, following the filing date of the parent nonprovisional application. An example of a proper benefit claim of a prior national stage of an international application is "This application is a continuation of U.S. Application No. X, which is the National Stage of International Application No. PCT/US - / - , filed - ." For additional examples of proper benefit claims, see Manual of Patent Examining Procedure (8th ed., August 2001)(MPEP), Section 201.11, Reference to First Application. Section 201.11 of the MPEP will be revised in the upcoming revision to reflect the clarification made in this notice about the required manner of making proper claims for the benefit of prior nonprovisional and provisional applications.

As stated previously, to specify the relationship between the nonprovisional applications, applicant must specify whether the subject application is a continuation, divisional, or continuation-in-part of the prior nonprovisional application. Note that the terms are exclusive. An application cannot be, for example, both a continuation

and a divisional, or a continuation and a continuation-in-part, of the same parent application. Moreover, if the benefit of more than one nonprovisional parent application is claimed, the relationship must include an identification of each nonprovisional application as a continuation, divisional, or continuation-in-part application of the immediate prior nonprovisional application for which a benefit is claimed in order to establish co-pendency throughout the entire chain of prior-filed parent nonprovisional applications. For example, the following two statements are improper: "This application claims the benefit of Application Nos. C, B, and A." and "This application is a continuing application of Application Nos. C, B, and A." On the other hand, the following statement is proper and acceptable: "This application is a continuation of Application No. C, filed - , which is a continuation of Application No. B, filed - , which is a continuation of Application No. A, filed - ."

Sanctions for Making Improper Benefit Claims to Nonprovisional Applications

Any benefit claim under 35 U.S.C. 120, 121 or 365(c) that does not identify a prior application and also specify a relationship between each of the applications will not be in compliance with 37 CFR 1.78(a)(2)(i), and will not be considered to contain a specific reference to a prior application as required by 35 U.S.C. 120. Such a benefit claim will not be recognized by the Office and will not be included on the filing receipt for the application, even if the claim appears in the first sentence of the specification or an application data sheet (37 CFR 1.76), because the Office does not have sufficient information to enter the benefit claim into the Office's database. As a result, publication of the application will not be scheduled as a function of the prior application's filing date. The Office plans to notify applicants on, or with, the filing receipt that a benefit claim may not have been recognized because it did not include the proper reference. Applicants are advised that only the benefit claims that are listed on the filing receipt have been recognized by the Office. Since the filing receipt and the notification will usually be provided to the applicant shortly after the filing of the application, applicants should have sufficient opportunity to submit the proper benefit claims within the time period set in 37 CFR 1.78(a)(2)(ii) and thus avoid the need to submit a petition under 37 CFR 1.78(a)(3) and the surcharge set forth in 37 CFR 1.17(t). Failure to timely submit the reference required by 37 CFR 1.78(a)(2)(i) is considered a waiver of any benefit claim under 35 U.S.C. 120, 121 or 365(c) unless a petition to accept an unintentionally delayed claim under 37 CFR 1.78(a)(3), the surcharge set forth in 37 CFR 1.17(t), and the required reference, including the relationship of the applications (unless previously submitted) are filed. For example, if a benefit claim is submitted without the specific relationship between the nonprovisional applications before the expiration of the period, and the specific relationship between the nonprovisional applications is subsequently submitted after the expiration of the period, a petition and the surcharge would be required.

Part II: Reference to Prior Provisional Application(s) Per 37 CFR 1.78(a)(5)(i) Should Only Include Identification of Prior Provisional Application(s)

When the domestic benefit of a prior provisional application is being claimed under 35 U.S.C. 119(e), however, the relationship between the two applications should not be specified. 35

U.S.C. 119(e) provides that a nonprovisional application claiming the benefit of a provisional application must be filed within twelve months of the provisional application and must contain, or be amended to contain, a specific reference to the provisional application. The specific reference required by 35 U.S.C. 119(e) is the reference required by 37 CFR 1.78(a)(5). 37 CFR 1.78(a)(5)(i) requires that any nonprovisional application, or international application designating the United States, claiming the benefit of one or more prior-filed provisional applications must contain, or be amended to contain, a reference to each such prior-filed provisional application identifying it by provisional application number. No relationship should be specified whenever a claim for the benefit of a provisional application under 35 U.S.C. 119(e) is made.

If a relationship between a nonprovisional application and a prior provisional application is submitted, however, it may be unclear whether applicant wishes to claim the domestic benefit of the provisional application under 35 U.S.C. 119(e), or the benefit of an earlier application's filing date under 35 U.S.C. 120. Thus, applicants seeking to claim the domestic benefit of a provisional application under 35 U.S.C. 119(e) should not state that the application is a "continuation" of a provisional application, nor should it be stated that the application claims benefit under 35 U.S.C. 120 of a provisional application. If such a claim is submitted in an application transmitted to the Office other than through the Electronic Filing System, it will be entered into the Office computer system as a claim to the "benefit" of the provisional application. Although 35 U.S.C. 120 does not preclude a benefit claim to a provisional application (that is, one could obtain the benefit under 35 U.S.C. 120 of a prior filed provisional application), such a benefit claim under 35 U.S.C. 120 is not recommended as such a claim may have the effect of reducing the patent term, as the term of a patent issuing from such an application may be measured from the filing date of the provisional application pursuant to 35 U.S.C. 154(a)(2). Instead, applicants should state "This application claims the benefit of U.S. Provisional Application No. 60/ - , filed - ", or "This application claims the benefit of U.S. Provisional Application No. 60/ - , filed - , and U.S. Provisional Application 60/ - , filed - ." See MPEP 201.11, Reference to First Application (8th ed., August 2001).

Part III: If Benefit is Claimed of a Prior Provisional Application Which was Filed More Than One Year Before the Subject Application, Then Each Prior Nonprovisional Application(s) Claiming Benefit of the Provisional Must be Specified

Any nonprovisional application that directly claims the benefit of a provisional application under 35 U.S.C. 119(e) must be filed within twelve months from the filing date of the provisional application. As noted above, an application that itself directly claims the benefit of a provisional application should identify, but not specify the relationship to, the provisional application. If the subject nonprovisional application is not filed within the twelve month period, however, it still may claim the benefit of the provisional application via an intermediate nonprovisional application under 35 U.S.C. 120. The intermediate nonprovisional application must have been filed within twelve months from the filing date of the provisional application and the intermediate nonprovisional application must have claimed the benefit of the provisional application. Further, it must be clearly indicated that the intermediate nonprovisional application is

claiming the benefit under 35 U.S.C. 119(e) of the provisional application. This identification of the intermediate nonprovisional application is necessary so that the Office can determine whether the intermediate nonprovisional application was filed within twelve months of the filing date of the provisional application, and thus, whether the benefit claim is proper.

Examples

Applicant should state such a benefit claim as follows: "This application is a continuation of Application No. C, filed - , which is a continuation of Application No. B, filed - , which claims the benefit of U.S. Provisional Application No. A, filed - ." A benefit claim that merely states "This application claims the benefit of nonprovisional Application No. C, filed - , nonprovisional Application No. B, filed - , and provisional application No. A, filed - " would be improper where the subject application was not filed within twelve months of the provisional application.

Where the benefit of more than one provisional application is being claimed, the intermediate nonprovisional application(s) claiming the benefit of each provisional application must be clearly indicated. Applicant should state, for example, "This application is continuation of Application No. D, filed - , which is a continuation-in-part of Application No. C, filed - , Application No. D claims the benefit of U.S. Provisional Application No. B, filed - , and Application No. C claims the benefit of U.S. Provisional Application No. A, filed - ." An example of a proper benefit claim of a prior national stage of an international application, which claims the priority to a provisional application, is "This application is a continuation of U.S. Application No. Y, which is the National Stage of International Application No. PCT/US - / - , filed - , which claims the benefit under 35 U.S.C. 119(e) of U.S. Provisional Application X, filed - ."

Sanctions for Making Improper Benefit Claims to Provisional Applications

If a benefit claim to a provisional application is submitted without an indication that an intermediate nonprovisional application directly claims the benefit of the provisional application and the instant nonprovisional application is not filed within the twelve month period, or the relationship between nonprovisional applications is not indicated, the Office will not have sufficient information to enter the benefit claim into the computer database. Therefore, the Office will not recognize such a benefit claim, and will not include the benefit claim on the filing receipt. The Office plans to notify applicants on, or with, the filing receipt that a benefit claim may not have been recognized because information regarding the intermediate nonprovisional application(s) and/or the relationship between each nonprovisional application have not been provided. Applicants are advised that only the benefit claims that are listed on the filing receipt have been recognized by the Office. Since the filing receipt and the notification will usually be provided to the applicant shortly after the filing of the application, applicants should have sufficient opportunity to submit the proper benefit claims within the time period set in 37 CFR 1.78(a) and thus avoid the need to submit a petition under 37 CFR 1.78(a) and the surcharge set forth in 37 CFR 1.17(t). Failure to timely submit the reference required by 37 CFR 1.78(a) is considered a waiver of any benefit claim under 35 U.S.C. 119(e),

120, 121 or 365(c) unless a petition under 37 CFR 1.78(a), the surcharge set forth in 37 CFR 1.17(t), identification of the intermediate nonprovisional application which claims the benefit to the provisional application, and the relationship between each nonprovisional application are filed.

Part IV: Office Practice to Not Require Petition and Surcharge if Benefit Claim is Not Present in the Proper Place But is Recognized By Office Continues But Applicants Are Advised That Proper Reference Must be Presented

The reference required by 37 CFR 1.78(a)(2) or (a)(5) must be included in an application data sheet (37 CFR 1.76), or the specification must contain, or be amended to contain, such reference in the first sentence following the title. Previously, the Office indicated that if an applicant includes a benefit claim in the application but not in the manner specified by 37 CFR 1.78(a) (e.g., if the claim is included in an oath or declaration or the application transmittal letter) within the time period set forth in 37 CFR 1.78(a), the Office will not require a petition under 37 CFR 1.78(a) and the surcharge under 37 CFR 1.17(t) to correct the claim if the information concerning the claim was recognized by the Office as shown by its inclusion on the filing receipt. If, however, a claim is included elsewhere in the application and not recognized by the Office as shown by its absence on the filing receipt, the Office will require a petition and the surcharge to correct the claim. See Requirements for Claiming the Benefit of Prior-Filed Applications Under Eighteen-Month Publication Provisions, 66 Fed. Reg. 67087, 67089-90 (Dec. 28, 2001). The Office will continue to follow this practice.

Sanctions for Making Improper Benefit Claims

Applicants are simply being advised by this notice that the Office will not recognize any benefit claim where there is no indication of the relationship between the nonprovisional applications, or no indication of the intermediate nonprovisional application that is directly claiming the benefit of a provisional application. Applicants are also reminded that, even if the Office has recognized a benefit claim that includes the proper reference by entering it into the Office's database and including it on applicant's filing receipt, the benefit claim is not a proper benefit claim under 35 U.S.C. 119(e) and/or 35 U.S.C. 120, and 37 CFR 1.78, unless the reference is included in an application data sheet, or the first sentence of the specification, and all other requirements are met.

Part V: Correcting or Adding a Benefit Claim After Filing

The Office will not grant a request for a corrected filing receipt to include a benefit claim unless a proper reference to the prior application(s) is included in the first sentence of the specification, or an application data sheet, within the time period required by 37 CFR 1.78(a). Any request for corrected filing receipt to include a corrected or added benefit claim must be submitted within the time period required by 37 CFR 1.78(a) and be accompanied by an amendment to the specification, or an application data sheet. If the proper reference was previously submitted, a copy of the amendment, the first page of the specification, or the application data sheet, containing the claim should be included with the request for corrected filing receipt. The Office plans to notify applicants on, or with, the

filing receipt that a benefit claim may not have been recognized because it did not include the proper reference. Applicants are advised that only the benefit claims that are listed on the filing receipt have been recognized by the Office. Since the filing receipt and the notification will usually be provided shortly after the filing of the application, applicants should have sufficient opportunity to submit the proper benefit claims within the time period set in 37 CFR 1.78(a) and thus avoid the need to submit a petition under 37 CFR 1.78(a) and the surcharge set forth in 37 CFR 1.17(t). Therefore, applicants should carefully and promptly review their filing receipts in order to avoid the need for a petition and the surcharge.

When an unintentionally delayed benefit claim is submitted with a petition under 37 CFR 1.78(a) and the surcharge set forth in 37 CFR 1.17(t), the benefit claim must include a proper reference to the prior application(s) in order for the petition to be granted. The reference to the prior application(s) must include: (1) the relationship between nonprovisional applications (i.e., continuation, divisional, or continuation-in-part), and (2) the indication of any intermediate application that is directly claiming the benefit of a provisional application, in order to establish copendency throughout the entire chain of prior applications.

Applicants are also reminded that, if an amendment to the specification, or an application data sheet (ADS), is submitted in an application under final rejection, the amendment or ADS must be in compliance with 37 CFR 1.116. The amendment or ADS filed in an application under final rejection will not be entered as a matter of right. See MPEP 714.12 and 714.13. Therefore, applicants should consider filing a request for continued examination (RCE) (including fee and submission) under 37 CFR 1.114 with the petition to accept an unintentionally delayed benefit claim, the surcharge, and an amendment that adds the proper reference to the first sentence of the specification or an ADS.

Part VI: Each Intermediate Prior Application Must Have Proper Reference

If the benefit of more than one prior application is claimed, applicant should also make sure that the proper references are made in each intermediate nonprovisional application in the chain of prior applications. If an applicant desires, for example, the following benefit claim: "This application is a continuation of Application No. C, filed - , which is a continuation of Application No. B, filed - , which claims the benefit of U.S. Provisional Application No. A, filed - ," then Application No. C must include a benefit claim containing a reference to Application No. B and provisional Application No. A, and Application No. B must include a benefit claim containing a reference to provisional Application No. A.

Part VII: Adding an Incorporation-By-Reference Statement in a Benefit Claim is Not Permitted After Filing

An incorporation-by-reference statement added after the filing date of an application is not permitted because no new matter can be added to an application after its filing date. See 35 U.S.C. 132(a). If an incorporation-by-reference statement is included in an amendment to the specification to add a benefit claim after the filing date of the application, the amendment would not be proper. When a benefit claim is submitted after the filing of an application, the reference to the prior application cannot include an incorporation-by-reference statement of the prior application. See Dart Industries v. Banner, 636 F.2d 684,

207 USPQ 273 (C.A.D.C. 1980). Therefore, the Office will not grant a petition to accept a benefit claim that includes an incorporation-by-reference statement of a prior application, unless the incorporation-by-reference statement was submitted on filing of the application.

Inquiries regarding this notice should be directed to Eugenia A. Jones or Joni Y. Chang, Legal Advisors, Office of Patent Legal Administration, by telephone at (703) 305-1622.

February 24, 2003

STEPHEN G. KUNIN
Deputy Commissioner for
Patent Examination Policy

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	)	
Kevin P. Baker et al.	)	
Serial No. 09/944,396	)	Examiner: Kemmerer, E.
Filing Date: August 30, 2001	)	Group Art Unit No.: 1646
For SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME	)	

DECLARATION OF AUDREY D. GODDARD, Ph.D UNDER 37 C.F.R. § 1.132

Assistant Commissioner of Patents
Washington, D.C. 20231

Sir:

I, Audrey D. Goddard, Ph.D. do hereby declare and say as follows:

1. I am a Senior Clinical Scientist at the Experimental Medicine/BioOncology, Medical Affairs Department of Genentech, Inc., South San Francisco, California 94080.

2. Between 1993 and 2001, I headed the DNA Sequencing Laboratory at the Molecular Biology Department of Genentech, Inc. During this time, my responsibilities included the identification and characterization of genes contributing to the oncogenic process, and determination of the chromosomal localization of novel genes.

3. My scientific Curriculum Vitae, including my list of publications, is attached to and forms part of this Declaration (Exhibit A).

4. I am familiar with a variety of techniques known in the art for detecting and quantifying the amplification of oncogenes in cancer, including the quantitative TaqMan PCR (i.e., "gene amplification") assay described in the above captioned patent application.

5. The TaqMan PCR assay is described, for example, in the following scientific publications: Higuchi *et al.*, Biotechnology 10:413-417 (1992) (Exhibit B); Livak *et al.*, PCR Methods Appl., 4:357-362 (1995) (Exhibit C) and Heid *et al.*, Genome Res. 6:986-994 (1996) (Exhibit D). Briefly, the assay is based on the principle that successful PCR yields a fluorescent signal due to Taq DNA polymerase-mediated exonuclease digestion of a fluorescently labeled oligonucleotide that is homologous to a sequence between two PCR primers. The extent of digestion depends directly on the amount of PCR, and can be quantified accurately by measuring the increment in fluorescence that results from decreased energy transfer. This is an extremely sensitive technique, which allows detection in the exponential phase of the PCR reaction and, as a result, leads to accurate determination of gene copy number.

6. The quantitative fluorescent TaqMan PCR assay has been extensively and successfully used to characterize genes involved in cancer development and progression. Amplification of protooncogenes has been studied in a variety of human tumors, and is widely considered as having etiological, diagnostic and prognostic significance. This use of the quantitative TaqMan PCR assay is exemplified by the following scientific publications: Pennica *et al.*, Proc. Natl. Acad. Sci. USA 95(25):14717-14722 (1998) (Exhibit E); Pitti *et al.*, Nature 396(6712):699-703 (1998) (Exhibit F) and Bieche *et al.*, Int. J. Cancer 78:661-666 (1998) (Exhibit G), the first two of which I am co-author. In particular, Pennica *et al.* have used the quantitative TaqMan PCR assay to study relative gene amplification of WISP and c-myc in various cell lines, colorectal tumors and normal mucosa. Pitti *et al.* studied the genomic amplification of a decoy receptor for Fas ligand in lung and colon cancer, using the quantitative TaqMan PCR assay. Bieche *et al.* used the assay to study gene amplification in breast cancer.

7. It is my personal experience that the quantitative TaqMan PCR technique is technically sensitive enough to detect at least a 2-fold increase in gene copy number relative to control. It is further my considered scientific opinion that an at least 2-fold increase in gene copy number in a tumor tissue sample relative to a normal (i.e., non-tumor) sample is significant and useful in that the detected increase in gene copy number in the tumor sample relative to the normal sample serves as a basis for using relative gene copy number as quantitated by the TaqMan PCR technique as a diagnostic marker for the presence or absence of tumor in a tissue sample of unknown pathology. Accordingly, a gene identified as being amplified at least 2-fold by the quantitative TaqMan PCR assay in a tumor sample relative to a normal sample is useful as a marker for the diagnosis of cancer, for monitoring cancer development and/or for measuring the efficacy of cancer therapy.

8. I declare further that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true. I declare that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Jan. 16, 2003
Date

Audrey Goddard
Audrey D. Goddard, Ph.D.

AUDREY D. GODDARD, Ph.D.

Genentech, Inc.
1 DNA Way
South San Francisco, CA, 94080
650.225.6429
goddarda@gene.com

110 Congo St.
San Francisco, CA, 94131
415.841.9154
415.819.2247 (mobile)
agoddard@pacbell.net

PROFESSIONAL EXPERIENCE

1993-present

Genentech, Inc.
South San Francisco, CA

2001 - present **Senior Clinical Scientist**
Experimental Medicine / BioOncology, Medical Affairs

Responsibilities:
- *Companion diagnostic oncology products*
- *Acquisition of clinical samples from Genentech's clinical trials for translational research*
- *Translational research using clinical specimen and data for drug development and diagnostics*
- *Member of Development Science Review Committee, Diagnostic Oversight Team, 21 CFR Part 11 Subteam*

Interests:
- *Ethical and legal implications of experiments with clinical specimens and data*
- *Application of pharmacogenomics in clinical trials*

1998 - 2001 **Senior Scientist**
Head of the DNA Sequencing Laboratory, Molecular Biology Department, Research

Responsibilities:
- *Management of a laboratory of up to nineteen –including postdoctoral fellow, associate scientist, senior research associate and research assistants/associate levels*
- *Management of a $750K budget*
- *DNA sequencing core facility supporting a 350+ person research facility.*
- *DNA sequencing for high throughput gene discovery. - ESTs, cDNAs, and constructs*
- *Genomic sequence analysis and gene identification*
- *DNA sequence and primary protein analysis*

Research:
- *Chromosomal localization of novel genes*
- *Identification and characterization of genes contributing to the oncogenic process*
- *Identification and characterization of genes contributing to inflammatory diseases*
- *Design and development of schemes for high throughput genomic DNA sequence analysis*
- *Candidate gene prediction and evaluation*

1993 - 1998 Scientist
Head of the DNA Sequencing Laboratory, Molecular Biology Department, Research

Responsibilities
- *DNA sequencing core facility supporting a 350+ person research facility*
- *Assumed responsibility for a pre-existing team of five technicians and expanded the group into fifteen, introducing a level of middle management and additional areas of research*
- *Participated in the development of the basic plan for high throughput secreted protein discovery program – sequencing strategies, data analysis and tracking, database design*
- *High throughput EST and cDNA sequencing for new gene identification.*
- *Design and implementation of analysis tools required for high throughput gene identification.*
- *Chromosomal localization of genes encoding novel secreted proteins.*

Research:
- *Genomic sequence scanning for new gene discovery.*
- *Development of signal peptide selection methods.*
- *Evaluation of candidate disease genes.*
- *Growth hormone receptor gene SNPs in children with Idiopathic short stature*

Imperial Cancer Research Fund **1989-1992**
London, UK with Dr. Ellen Solomon

6/89 –12/92 Postdoctoral Fellow
- Cloning and characterization of the genes fused at the acute promyelocytic leukemia translocation breakpoints on chromosomes 17 and 15.
- Prepared a successfully funded European Union multi-center grant application

McMaster University **1983**
Hamilton, Ontario, Canada with Dr. G. D. Sweeney

5/83 – 8/83: NSERC Summer Student
- *In vitro* metabolism of β-naphthoflavone in C57Bl/6J and DBA mice

EDUCATION

Ph.D. "Phenotypic and genotypic effects of mutations in the human retinoblastoma gene." **Supervisor:** Dr. R. A. Phillips	University of Toronto Toronto, Ontario, Canada. Department of Medical Biophysics.	1989
Honours B.Sc "The *In vitro* metabolism of the cytochrome P-448 inducer β-naphthoflavone in C57BL/6J mice." **Supervisor:** Dr. G. D. Sweeney	McMaster University, Hamilton, Ontario, Canada. Department of Biochemistry	1983

ACADEMIC AWARDS

Imperial Cancer Research Fund Postdoctoral Fellowship	1989-1992
Medical Research Council Studentship	1983-1988
NSERC Undergraduate Summer Research Award	1983
Society of Chemical Industry Merit Award (Hons. Biochem.)	1983
Dr. Harry Lyman Hooker Scholarship	1981-1983
J.L.W. Gill Scholarship	1981-1982
Business and Professional Women's Club Scholarship	1980-1981
Wyerhauser Foundation Scholarship	1979-1980

INVITED PRESENTATIONS

Genentech's gene discovery pipeline: High throughput identification, cloning and characterization of novel genes. Functional Genomics: From Genome to Function, Litchfield Park, AZ, USA. October 2000

High throughput identification, cloning and characterization of novel genes. G2K:Back to Science, Advances in Genome Biology and Technology I. Marco Island, FL, USA. February 2000

Quality control in DNA Sequencing: The use of Phred and Phrap. Bay Area Sequencing Users Meeting, Berkeley, CA, USA. April 1999

High throughput secreted protein identification and cloning. Tenth International Genome Sequencing and Analysis Conference, Miami, FL, USA. September 1998

The evolution of DNA sequencing: The Genentech perspective. Bay Area Sequencing Users Meeting, Berkeley, CA, USA. May 1998

Partial Growth Hormone Insensitivity: The role of GH-receptor mutations in Idiopathic Short Stature. Tenth Annual National Cooperative Growth Study Investigators Meeting, San Francisco, CA, USA. October, 1996

Growth hormone (GH) receptor defects are present in selected children with non-GH-deficient short stature: A molecular basis for partial GH-insensitivity. 76th Annual Meeting of The Endocrine Society, Anaheim, CA, USA. June 1994

A previously uncharacterized gene, myl, is fused to the retinoic acid receptor alpha gene in acute promyelocytic leukemia. XV International Association for Comparative Research on Leukemia and Related Disease, Padua, Italy. October 1991

PATENTS

Goddard A, Godowski PJ, Gurney AL. NL2 Tie ligand homologue polypeptide. Patent Number: 6,455,496. Date of Patent: Sept. 24, 2002.

Goddard A, Godowski PJ and Gurney AL. NL3 Tie ligand homolngue nucleic acids. Patent Number: 6,426,218. Date of Patent: July 30, 2002.

Godowski P, Gurney A, Hillan KJ, Botstein D, **Goddard A**, Roy M, Ferrara N, Tumas D, Schwall R. NL4 Tie ligand homologue nucleic acid. Patent Number: 6,4137,770. Date of Patent: July 2, 2002.

Ashkenazi A, Fong S, **Goddard A**, Gurney AL, Napier MA, Tumas D, Wood WI. Nucleic acid encoding A-33 related antigen poly peptides. Patent Number: 6,410,708. Date of Patent:: Jun. 25, 2002.

Botstein DA, Cohen RL, **Goddard AD**, Gurney AL, Hillan KJ, Lawrence DA, Levine AJ, Pennica D, Roy MA and Wood WI. WISP polypeptides and nucleic acids encoding same. Patent Number: 6,387,657. Date nf Patent: May 14, 2002.

Goddard A, Godowski PJ and Gurney AL. Tie ligands. Patent Number: 6,372,491. Date of Patent: April 16, 2002.

Godowski PJ, Gurney AL, **Goddard A** and Hillan K. TIE ligand homolngue antibody. Patent Number: 6,350,450. Date of Patent: Feb. 26, 2002.

Fong S, Ferrara N, **Goddard** A, Godowski PJ, Gurney AL, Hillan K and Williams PM. Tie receptor tyrosine kinase ligand homologues. Patent Number: 6,348,351. Date of Patent: Feb. 10, 2002.

Goddard A, Godowski PJ and Gurney AL. Ligand homologues. Patent Number: 6,348,350. Date of Patent: Feb. 19, 2002.

Attie KM, Carlsson LMS, Gesundheit N and **Goddard A**. Treatment of partial growth hormone Insensitivity syndrome. Patent Number: 6,207,640. Date of Patent: March 27, 2001.

Fong S, Ferrara N, **Goddard A**, Godowski PJ, Gurney AL, Hillan K and Williams PM. Nucleic acids encoding NL-3. Patent Number: 6,074,673. Date of Patent: June 13, 2000

Attie K, Carlsson LMS, Gesunheit N and **Goddard A**. Treatment of partial growth hormone Insensitivity syndrome. Patent Number: 5,824,642. Date of Patent: October 20, 1998

Attie K, Carlsson LMS, Gesunheit N and **Goddard A**. Treatment of partial growth hormone Insensitivity syndrome. Patent Number: 5,646,113. Date of Patent: July 8, 1997

Multiple additional provisional applications filed

PUBLICATIONS

Seshasayee D, Dowd P, Gu Q, Erickson S, **Goddard AD** Comparative sequence analysis of the *HER2* locus in mouse and man. Manuscript in preparation.

Abuzzahab MJ, **Goddard A**, Grigorescu F, Lautier C, Smith RJ and Chernausek SD. Human IGF-1 receptor mutations resulting in pre- and post-natal growth retardation. Manuscript in preparation.

Aggarwal S, Xie, M-H, Foster J, Frantz G, Stinson J, Corpuz RT, Simmons L, Hillan K, Yansura DG, Vandlen RL, **Goddard AD** and Gurney AL. FHFR, a novel receptor for the fibroblast growth factors. Manuscript submitted.

Adams SH, Chui C, Schilbach SL, Yu XX, **Goddard AD**, Grimaldi JC, Lee J, Dowd P, Colman S., Lewin DA. (2001) BFIT, a unique acyl-CoA thioesterase induced in thermogenic brown adipose tissue: Cloning, organization of the human gene, and assessment of a potential link to obesity. *Biochemical Journal* **360**: 135-142.

Lee J, Ho WH, Maruoka M, Corpuz RT, Baldwin DT, Foster JS, **Goddard AD**, Yansura DG, Vandlen RL, Wood WI, Gurney AL. (2001) IL-17E, a novel proinflammatory ligand for the IL-17 receptor homolog IL-17Rh1. *Journal of Biological Chemistry* **276**(2): 1660-1664.

Xie M-H, Aggarwal S, Ho W-H, Foster J, Zhang Z, Stinson J, Wood WI, **Goddard AD** and Gurney AL. (2000) Interleukin (IL)-22, a novel human cytokine that signals through the interferon-receptor related proteins CRF2-4 and IL-22R. *Journal of Biological Chemistry* **275**: 31335-31339.

Weiss GA, Watanabe CK, Zhong A, **Goddard A** and Sidhu SS. (2000) Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proc. Natl. Acad. Sci. USA* **97**: 8050-8954.

Guo S, Yamaguchi Y, Schilbach S, Wada T, Lee J, **Goddard A**, French D, Handa H, Rosenthal A. (2000) A regulator of transcriptional elongation controls vertebrate neuronal development. *Nature* **408**: 366-369.

Yan M, Wang L-C, Hymowitz SG, Schilbach S, Lee J, **Goddard A**, de Vos AM, Gao WQ, Dixit VM. (2000) Two-amino acid molecular switch in an epithelial morphogen that regulates binding to two distinct receptors. *Science* **290**: 523-527.

Sehl PD, Tai JTN, Hillan KJ, Brown LA, **Goddard A**, Yang R, Jin H and Lowe DG. (2000) Application of cDNA microarrays in determining molecular phenotype in cardiac growth, development, and response to injury. *Circulation* **101**: 1990-1999.

Guo S, Brush J, Teraoka H, **Goddard A**, Wilson SW, Mullins MC and Rosenthal A. (1999) Development of noradrenergic neurons in the zebrafish hindbrain requires BMP, FGF8, and the homeodomain protein soulless/Phox2A. *Neuron* **24**: 555-566.

Stone D, Murone, M, Luoh, S, Ye W, Armanini P, Gurney A, Phillips HS, Brush, J, **Goddard A**, de Sauvage FJ and Rosenthal A. (1999) Characterization of the human suppressor of fused: a negative regulator of the zinc-finger transcription factor Gli. *J. Cell Sci.* **112**: 4437-4448.

Xie M-H, Holcomb I, Deuel B, Dowd P, Huang A, Vagts A, Foster J, Liang J, Brush J, Gu Q, Hillan K, **Goddard A** and Gurney, A.L. (1999) FGF-19, a novel fibroblast growth factor with unique specificity for FGFR4. *Cytokine* **11**: 729-735.

Yan M, Lee J, Schilbach S, **Goddard A** and Dixit V. (1999) mE10, a novel caspase recruitment domain-contalning proapoptotic molecule. *J. Biol. Chem.* 274(15): 10287-10292.

Gurney AL, Marsters SA, Huang RM, Pitti RM, Mark DT, Baldwin DT, Gray AM, Dowd P, Brush J, Heldens S, Schow P, **Goddard AD**, Wood WI, Baker KP, Godowski PJ and Ashkenazi A. (1999) Identification of a new member of the tumor necrosis factor family and its receptor, a human ortholog of mouse GITR. *Current Biology* 9(4): 215-218.

Ridgway JBB, Ng E, Kern JA ,Lee J, Brush J, **Goddard A** and Carter P. (1999) Identification of a human anti-CD55 single-chain Fv by subtractive panning of a phage library using tumor and nontumor cell lines. *Cancer Research* 59: 2718-2723.

Pitti RM, Marsters SA, Lawrence DA, Roy M, Kischkel FC, Dowd P, Huang A, Donahue CJ, Sherwood SW, Baldwin DT, Godowski PJ, Wood WI, Gurney AL, Hillan KJ, Cohen RL, **Goddard AD**, Botstein D and Ashkenazi A. (1998) Genomic amplification of a decoy receptor for Fas ligand in lung and colon cancer. *Nature* 396(6712): 699-703.

Pennica D, Swanson TA, Welsh JW, Roy MA, Lawrence DA, Lee J, Brush J, Taneyhill LA, Deuel B, Lew M, Watanabe C, Cohen RL, Melhem MF, Finley GG, Quirke P, **Goddard AD**, Hillan KJ, Gurney AL, Botstein D and Levine AJ. (1998) WISP genes are members of the connective tissue growth factor family that are up-regulated in wnt-1-transformed cells and aberrantly expressed in human colon tumors. *Proc. Natl. Acad. Sci. USA.* 95(25): 14717-14722.

Yang RB, Mark MR, Gray A, Huang A, Xie MH, Zhang M, **Goddard A**, Wood WI, Gurney AL and Godowski PJ. (1998) Toll-like receptor-2 mediates lipopolysaccharide-induced cellular signalling. *Nature* 395(6699): 284-288.

Merchant AM, Zhu Z, Yuan JQ, **Goddard A**, Adams CW, Presta LG and Carter P. (1998) An efficient route to human bispecific IgG. *Nature Biotechnology* 16(7): 677-681.

Marsters SA, Sheridan JP, Pitti RM, Brush J, **Goddard A** and Ashkenazi A. (1998) Identification of a ligand for the death-domain-containing receptor Apo3. *Current Biology* 8(9): 525-528.

Xie J, Murone M, Luoh SM, Ryan A, Gu Q, Zhang C, Bonifas JM, Lam CW, Hynes M, **Goddard A**, Rosenthal A, Epstein EH Jr. and de Sauvage FJ. (1998) Activating Smoothened mutations in sporadic basal-cell carcinoma. *Nature*. 391(6662): 90-92.

Marsters SA, Sheridan JP, Pitti RM, Huang A, Skubatch M, Baldwin D, Yuan J, Gurney A, **Goddard AD**, Godowski P and Ashkenazi A. (1997) A novel receptor for Apo2L/TRAIL contains a truncated death domain. *Current Biology*. 7(12): 1003-1006.

Hynes M, Stone DM, Dowd M, Pitts-Meek S, **Goddard A**, Gurney A and Rosenthal A. (1997) Control of cell pattern in the neural tube by the zinc finger transcription factor *Gli-1*. *Neuron* 19: 15–26.

Sheridan JP, Marsters SA, Pitti RM, Gurney A, Skubatch M, Baldwin D, Ramakrishnan L, Gray CL, Baker K, Wood WI, **Goddard AD**, Godowski P, and Ashkenazi A. (1997) Control of TRAIL-Induced Apoptosis by a Family of Signaling and Decoy Receptors. *Science* 277 (5327): 818-821.

Goddard AD, Dowd P, Chernausek S, Geffner M, Gertner J, Hintz R, Hopwood N, Kaplan S, Plotnick L, Rogol A, Rosenfield R, Saenger P, Mauras N, Hershkopf R, Angulo M and Attie, K. (1997) Partial growth hormone insensitivity: The role of growth hormone receptor mutations in idiopathic short stature. *J. Pediatr.* 131: S51-55.

Klein RD, Sherman D, Ho WH, Stone D, Bennett GL, Moffat B, Vandlen R, Simmons L, Gu Q, Hongo JA, Devaux B, Poulsen K, Armanini M, Nozaki C, Asai N, **Goddard A**, Phillips H, Henderson CE, Takahashi M and Rosenthal A. (1997) A GPI-linked protein that interacts with Ret to form a candidate neurturin receptor. *Nature*. 387(6634): 717-21.

Stone DM, Hynes M, Armanini M, Swanson TA, Gu Q, Johnson RL, Scott MP, Pennica D, **Goddard A**, Phillips H, Noll M, Hooper JE, de Sauvage F and Rosenthal A. (1996) The tumour-suppressor gene patched encodes a candidate receptor for Sonic hedgehog. *Nature* 384(6605): 129-34.

Marsters SA, Sheridan JP, Donahue CJ, Pitti RM, Gray CL, **Goddard AD**, Bauer KD and Ashkenazi A. (1996) Apo-3, a new member of the tumor necrosis factor receptor family, contains a death domain and activates apoptosis and NF-kappa B. *Current Biology* 6(12): 1669-76.

Rothe M, Xiong J, Shu HB, Williamson K, **Goddard A** and Goeddel DV. (1996) I-TRAF is a novel TRAF-interacting protein that regulates TRAF-mediated signal transduction. *Proc. Natl. Acad. Sci. USA* 93: 8241-8246.

Yang M, Luoh SM, **Goddard A**, Reilly D, Henzel W and Bass S. (1996) The bglX gene located at 47.8 min on the Escherichia coli chromosome encodes a periplasmic beta-glucosidase. *Microbiology* 142: 1659-65.

Goddard AD and Black DM. (1996) Familial Cancer In Molecular Endocrinology of Cancer. Waxman, J. Ed. Cambridge University Press, Cambridge UK, pp.187-215.

Treanor JJS, Goodman L, de Sauvage F, Stone DM, Poulson KT, Beck CD, Gray C, Armanini MP, Pollocks RA, Hefti F, Phillips HS, **Goddard A**, Moore MW, Buj-Bello A, Davis AM, Asai N, Takahashi M, Vandlen R, Henderson CE and Rosenthal A. (1996) Characterization of a receptor for GDNF. *Nature* 382: 80-83.

Klein RD, Gu Q, **Goddard A** and Rosenthal A. (1996) Selection for genes encoding secreted proteins and receptors. *Proc. Natl. Acad. Sci. USA* 93: 7108-7113.

Winslow JW, Moran P, Valverde J, Shih A, Yuan JQ, Wong SC, Tsai SP, **Goddard A**, Henzel WJ, Hefti F and Caras I. (1995) Cloning of AL-1, a ligand for an Eph-related tyrosine kinase receptor involved in axon bundle formation. *Neuron* 14: 973-981.

Bennett BD, Zeigler FC, Gu Q, Fendly B, **Goddard AD**, Gillett N and Matthews W. (1995) Molecular cloning of a ligand for the EPH-related receptor protein-tyrosine kinase Htk. *Proc. Natl. Acad. Sci. USA* 92: 1866-1870.

Huang X, Yuang J, **Goddard A**, Foulis A, James RF, Lemmark A, Pujol-Borrell R, Rabinovitch A, Somoza N and Stewart TA. (1995) Interferon expression in the pancreases of patients with type I diabetes. *Diabetes* 44: 658-664.

Goddard AD, Yuan JQ, Fairbairn L, Dexter M, Borrow J, Kozak C and Solomon E. (1995) Cloning of the murine homolog of the leukemia-associated FML gene. *Mammalian Genome* 6: 732-737.

Goddard AD, Covello R, Luoh SM, Clackson T, Attie KM, Gesundheit N, Rundle AC, Wells JA, Carlsson LMTI and The Growth Hormone Insensitivity Study Group. (1995) Mutations of the growth hormone receptor in children with idiopathic short stature. *N. Engl. J. Med.* **333**: 1093-1098.

Kuo SS, Moran P, Gripp J, Armanini M, Phillips HS, **Goddard A** and Caras IW. (1994) Identification and characterization of Batk, a predominantly brain-specific non-receptor protein tyrosine kinase related to Csk. *J. Neurosci. Res.* **38**: 705-715.

Mark MR, Scadden DT, Wang Z, Gu Q, **Goddard A** and Godowski PJ. (1994) Rse, a novel receptor-type tyrosine kinase with homology to Axl/Ufo, is expressed at high levels in the brain. *Journal of Biological Chemistry* **269**: 10720-10728.

Borrow J, Shipley J, Howe K, Kiely F, **Goddard A**, Sheer D, Srivastava A, Antony AC, Fioretos T, Mitelman F and Solomon E. (1994) Molecular analysis of simple variant translocations in acute promyelocytic leukemia. *Genes Chromosomes Cancer* **9**: 234-243.

Goddard AD and Solomon E. (1993) Genetics of Cancer. *Adv. Hum. Genet.* **21**: 321-376.

Borrow J, **Goddard AD**, Gibbons B, Katz F, Swirsky D, Floretos T, Dube I, Winfield DA, Kingston J, Hagemeijer A, Rees JKH, Lister AT and Solomon E. (1992) Diagnosis of acute promyelocytic leukemia by RT-PCR: Detection of *PML-RARA* and *RARA-PML* fusion transcripts. *Br. J. Haematol.* **82**: 529-540.

Goddard AD, Borrow J and Solomon E. (1992) A previously uncharacterized gene, PML, is fused to the retinoic acid receptor alpha gene in acute promyelocytic leukemia. *Leukemia* **6 Suppl 3**: 117S–119S.

Zhu X, Dunn JM, **Goddard AD**, Squire JA, Becker A, Phillips RA and Gallie BL. (1992) Mechanisms of loss of heterozygosity in retinoblastoma. *Cytogenet. Cell. Genet.* **59**: 248-252.

Foulkes W, **Goddard A** and Patel K. (1991) Retinoblastoma linked with Seascale [letter]. *British Med. J.* **302**: 409.

Goddard AD, Borrow J, Freemont PS and Solomon E. (1991) Characterization of a novel zinc finger gene disrupted by the t(15;17) in acute promyelocytic leukemia. *Science* **254**: 1371-1374.

Solomon E, Borrow J and **Goddard AD**. (1991) Chromosomal aberrations in cancer. *Science* **254**: 1153-1160.

Pajunen L, Jones TA, **Goddard A**, Sheer D, Solomon E, Pihlajaniemi T and Kivirikko KI. (1991) Regional assignment of the human gene coding for a multifunctional peptide (P4HB) acting as the β–subunit of prolyl-4-hydroxylase and the enzyme protein disulfide isomerase to 17q25. *Cytogenet. Cell. Genet.* **56**: 165-168.

Borrow J, Black DM, **Goddard AD**, Yagle MK, Frischauf A.-M and Solomon E. (1991) Construction and regional localization of a *Not*I linking library from human chromosome 17q. *Genomics* **10**: 477–480.

Borrow J, **Goddard AD**, Sheer D and Solomon E. (1990) Molecular analysis of acute promyelocytic leukemia breakpoint cluster region on chromosome 17. *Science* **249**: 1577-1580.

Myers JC, Jones TA, Pohjolainen E-R, Kadri AS, **Goddard AD**, Sheer D, Solomon E and Pihlajaniemi T. (1990) Molecular cloning of 5(IV) collagen and assignment of the gene to the region of the region of the X-chromosome containing the Alport Syndrome locus. *Am. J. Hum. Genet.* 46: 1024-1033.

Gallie BL, Squire JA, **Goddard A**, Dunn JM, Canton M, Hinton D, Zhu X and Phillips RA. (1990) Mechanisms of oncogenesis in retinoblastoma. *Lab. Invest.* 62: 394-408.

Goddard AD, Phillips RA, Greger V, Passarge E, Hopping W, Gallie BL and Horsthemke B. (1990) Use of the RB1 cDNA as a diagnostic probe in retinoblastoma families. *Clinical Genetics* 37: 117-126.

Zhu XP, Dunn JM, Phillips RA, **Goddard AD**, Paton KE, Becker A and Gallie BL. (1989) Germline, but not somatic, mutations of the RB1 gene preferentially involve the paternal allele. *Nature* 340: 312-314.

Gallie BL, Dunn JM, **Goddard A**, Becker A and Phillips RA. (1988) Identification of mutations in the putative retinoblastoma gene. In Molecular Biology of The Eye: Genes, Vision and Ocular Disease. UCLA Symposia on Molecular and Cellular Biology, New Series, Volume 88. J. Piatigorsky, T. Shinohara and P.S. Zelenka, Eds. Alan R. Liss, Inc., New York, 1988, pp. 427-436.

Goddard AD, Balakier H, Canton M, Dunn J, Squire J, Reyes E, Becker A, Phillips RA and Gallie BL. (1988) Infrequent genomic rearrangement and normal expression of the putative RB1 gene in retinoblastoma tumors. *Mol. Cell. Biol.* 8: 2082-2088.

Squire J, Dunn J, **Goddard A**, Hoffman T, Musarella M, Willard HF, Becker AJ, Gallie BL and Phillips RA. (1986) Cloning of the esterase D gene: A polymorphic gene probe closely linked to the retinoblastoma locus on chromosome 13. *Proc. Natl. Acad. Sci.* USA 83: 6573-6577.

Squire J, **Goddard AD**, Canton M, Becker A, Phillips RA and Gallie BL (1986) Tumour induction by the retinoblastoma mutation is independent of N-*myc* expression. *Nature* 322: 555-557.

Goddard AD, Heddle JA, Gallie BL and Phillips RA. (1985) Radiation sensitivity of fibroblasts of bilateral retinoblastoma patients as determined by micronucleus induction *in vitro*. *Mutation Research* 152: 31-38.

RESEARCH

SIMULTANEOUS AMPLIFICATION AND DETECTION OF SPECIFIC DNA SEQUENCES

Russell Higuchi*, Gavin Dollinger[1], P. Scan Walsh and Robert Griffith

Roche Molecular Systems, Inc., 1400 53rd St., Emeryville, CA 94608. [1]Chiron Corporation, 1400 53rd St., Emeryville, CA 94608. *Corresponding author.

We have enhanced the polymerase chain reaction (PCR) such that specific DNA sequences can be detected without opening the reaction tube. This enhancement requires the addition of ethidium bromide (EtBr) to a PCR. Since the fluorescence of EtBr increases in the presence of double-stranded (ds) DNA an increase in fluorescence in such a PCR indicates a positive amplification, which can be easily monitored externally. In fact, amplification can be continuously monitored in order to follow its progress. The ability to simultaneously amplify specific DNA sequences and detect the product of the amplification both simplifies and improves PCR and may facilitate its automation and more widespread use in the clinic or in other situations requiring high sample throughput.

Although the potential benefits of PCR[1] to clinical diagnostics are well known[2,3], it is still not widely used in this setting, even though it is four years since thermostable DNA polymerases made PCR practical. Some of the reasons for its slow acceptance are high cost, lack of automation of pre- and post-PCR processing steps, and false positive results from carryover-contamination. The first two points are related in that labor is the largest contributor to cost at the present stage of PCR development. Most current assays require some form of "downstream" processing once thermocycling is done in order to determine whether the target DNA sequence was present and has amplified. These include DNA hybridization[5,6], gel electrophoresis with or without use of restriction digestion[7,8], HPLC[9], or capillary electrophoresis[10]. These methods are labor-intense, have low throughput, and are difficult to automate. The third point is also closely related to downstream processing. The handling of the PCR product in these downstream processes increases the chances that amplified DNA will spread through the typing lab, resulting in a risk of "carryover" false positives in subsequent testing[11].

These downstream processing steps would be eliminated if specific amplification and detection of amplified DNA took place simultaneously within an unopened reaction vessel. Assays in which such different processes take place without the need to separate reaction components have been termed "homogeneous". No truly homogeneous PCR assay has been demonstrated to date, although progress towards this end has been reported. Chehab, et al.[12], developed a PCR product detection scheme using fluorescent primers that resulted in a fluorescent PCR product. Allele-specific primers, each with different fluorescent tags, were used to indicate the genotype of the DNA. However, the unincorporated primers must still be removed in a downstream process in order to visualize the result. Recently, Holland, et al.[13], developed an assay in which the endogenous 5' exonuclease assay of *Taq* DNA polymerase was exploited to cleave a labeled oligonucleotide probe. The probe would only cleave if PCR amplification had produced its complementary sequence. In order to detect the cleavage products, however, a subsequent process is again needed.

We have developed a truly homogeneous assay for PCR and PCR product detection based upon the greatly increased fluorescence that ethidium bromide and other DNA binding dyes exhibit when they are bound to dsDNA[14-16]. As outlined in Figure 1, a prototypic PCR

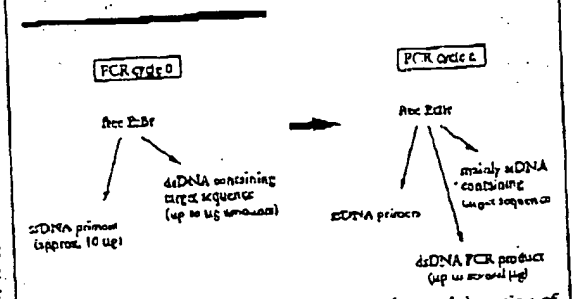


FIGURE 1 Principle of simultaneous amplification and detection of PCR product. The components of a PCR containing EtBr that are fluorescent are listed—EtBr itself, EtBr bound to either ssDNA or dsDNA. There is a large fluorescence enhancement when EtBr is bound to DNA and binding is greatly enhanced when DNA is double-stranded. After sufficient (n) cycles of PCR, the net increase in dsDNA results in additional EtBr binding, and a net increase in total fluorescence.

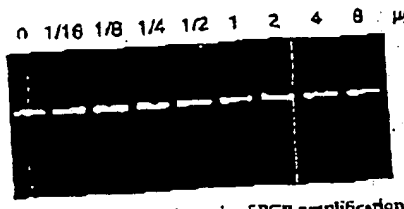


FIGURE 2 Gel electrophoresis of PCR amplification products of the human, nuclear gene, HLA DQα, made in the presence of increasing amounts of EtBr (up to 8 µg/ml). The presence of EtBr has no obvious effect on the yield or specificity of amplification.

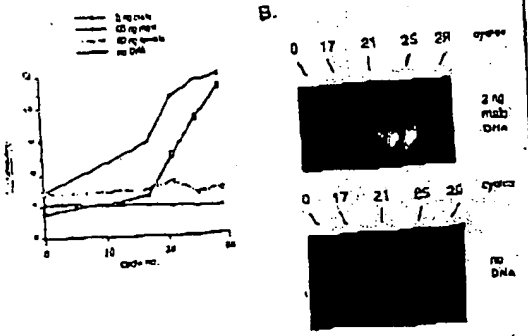


FIGURE 3 (A) Fluorescence measurements from PCRs that contain 0.5 µg/ml EtBr and that are specific for Y-chromosome repeat sequences. Five replicate PCRs were begun containing each of the DNAs specified. At each indicated cycle, one of the five replicate PCRs for each DNA was removed from thermocycling and its fluorescence measured. Units of fluorescence are arbitrary. (B) UV photography of PCR tubes (0.6 ml Eppendorf-style, polypropylene micro-centrifuge tubes) containing reactions, those starting from 2 ng male DNA and control reactions without any DNA, from (A).

begins with primers that are single-stranded DNA (ssDNA), dNTPs, and DNA polymerase. An amount of dsDNA containing the target sequence (target DNA) is also typically present. This amount can vary, depending on the application, from single-cell amounts of DNA[17] to micrograms per PCR[18]. If EtBr is present, the reagents that will fluoresce, in order of increasing fluorescence, are free EtBr itself, and EtBr bound to the single-stranded DNA primers and to the double-stranded target DNA (by its intercalation between the stacked bases of the DNA double-helix). After the first denaturation cycle, target DNA will be largely single-stranded. After a PCR is completed, the most significant change is the increase in the amount of dsDNA (the PCR product itself) of up to several micrograms. Formerly free EtBr is bound to the additional dsDNA, resulting in an increase in fluorescence. There is also some decrease in the amount of ssDNA primer, but because the binding of EtBr to ssDNA is much less than to dsDNA, the effect of this change on the total fluorescence of the sample is small. The fluorescence increase can be measured by directing excitation illumination through the walls of the amplification vessel before and after, or even continuously during, thermocycling.

RESULTS

PCR in the presence of EtBr. In order to assess the affect of EtBr in PCR, amplifications of the human HLA DQα gene[19] were performed with the dye present at concentrations from 0.06 to 8.0 µg/ml (a typical concentration of EtBr used in staining of nucleic acids following gel electrophoresis is 0.5 µg/ml). As shown in Figure 2, gel electrophoresis revealed little or no difference in the yield or quality of the amplification product whether EtBr was absent or present at any of these concentrations, indicating that EtBr does not inhibit PCR.

Detection of human Y-chromosome specific sequences. Sequence-specific, fluorescence enhancement of EtBr as a result of PCR was demonstrated in a series of amplifications containing 0.5 µg/ml EtBr and primers specific to repeat DNA sequences found on the human Y-chromosome[20]. These PCRs initially contained either 60 ng male, 60 ng female, 2 ng male human or no DNA. Five replicate PCRs were begun for each DNA. After 0, 17, 21, 24 and 29 cycles of thermocycling, a PCR for each DNA was removed from the thermocycler, and its fluorescence measured in a spectrofluorometer and plotted vs. amplification cycle number (Fig. 3A). The shape of this curve reflects the fact that by the time an increase in fluorescence can be detected, the increase in DNA is becoming linear and not exponential with cycle number. As shown, the fluorescence increased about three-fold over the background fluorescence for the PCRs containing human male DNA, but did not significantly increase for negative control PCRs, which contained either no DNA or human female DNA. The more male DNA present to begin with—60 ng versus 2 ng—the fewer cycles were needed to give a detectable increase in fluorescence. Gel electrophoresis on the products of these amplifications showed that DNA fragments of the expected size were made in the male DNA containing reactions and that little DNA synthesis took place in the control samples.

In addition, the increase in fluorescence was visualized by simply laying the completed, unopened PCRs on a UV transilluminator and photographing them through a red filter. This is shown in figure 3B for the reactions that began with 2 ng male DNA and those with no DNA.

Detection of specific alleles of the human β-globin gene. In order to demonstrate that this approach has adequate specificity to allow genetic screening, a detection of the sickle-cell anemia mutation was performed. Figure 4 shows the fluorescence from completed amplifications containing EtBr (0.5 µg/ml) as detected by photography of the reaction tubes on a UV transilluminator. These reactions were performed using primers specific for either the wild-type or sickle-cell mutation of the human β-globin gene[21]. The specificity for each allele is imparted by placing the sickle-mutation site at the terminal 3' nucleotide of one primer. By using an appropriate primer annealing temperature, primer extension—and thus amplification—can take place only if the 3' nucleotide of the primer is complementary to the β-globin allele present[21,22].

Each pair of amplifications shown in Figure 4 consists of a reaction with either the wild-type allele specific (left tube) or sickle-allele specific (right tube) primers. Three different DNAs were typed: DNA from a homozygous wild-type β-globin individual (AA); from a heterozygous sickle β-globin individual (AS); and from a homozygous sickle β-globin individual (SS). Each DNA (50 ng genomic DNA to start each PCR) was analyzed in triplicate (3 pairs

of reactions each). The DNA type was reflected in the relative fluorescence intensities in each pair of completed amplifications. There was a significant increase in fluorescence only where a β-globin allele DNA matched the primer set. When measured on a spectrofluorometer (data not shown), this fluorescence was about three times that present in a PCR where both β-globin alleles were mismatched to the primer set. Gel electrophoresis (not shown) established that this increase in fluorescence was due to the synthesis of nearly a microgram of a DNA fragment of the expected size for β-globin. There was little synthesis of dsDNA in reactions in which the allele-specific primer was mismatched to both alleles.

Continuous monitoring of a PCR. Using a fiber optic device, it is possible to direct excitation illumination from a spectrofluorometer to a PCR undergoing thermocycling and to return its fluorescence to the spectrofluorometer. The fluorescence readout of such an arrangement, directed at an EtBr-containing amplification of Y-chromosome specific sequences from 25 ag of human male DNA, is shown in Figure 5. The readout from a control PCR with no target DNA is also shown. Thirty cycles of PCR were monitored for each.

The fluorescence trace as a function of time clearly shows the effect of the thermocycling. Fluorescence intensity rises and falls inversely with temperature. The fluorescence intensity is minimum at the denaturation temperature (94°C) and maximum at the annealing/extension temperature (50°C). In the negative-control PCR, these fluorescence maxima and minima do not change significantly over the thirty thermocycles, indicating that there is little dsDNA synthesis without the appropriate target DNA, and there is little if any bleaching of EtBr during the continuous illumination of the sample.

In the PCR containing male DNA, the fluorescence maxima at the annealing/extension temperature begin to increase at about 4000 seconds of thermocycling, and continue to increase with time, indicating that dsDNA is being produced at a detectable level. Note that the fluorescence minima at the denaturation temperature do not significantly increase, presumably because at this temperature there is no dsDNA for EtBr to bind. Thus the course of the amplification is followed by tracking the fluorescence increase at the annealing temperature. Analysis of the products of these two amplifications by gel electrophoresis showed a DNA fragment of the expected size for the male DNA containing sample and no detectable DNA synthesis for the control sample.

DISCUSSION

Downstream processes such as hybridization to a sequence-specific probe can enhance the specificity of DNA detection by PCR. The elimination of these processes means that the specificity of this homogeneous assay depends solely on that of PCR. In the case of sickle-cell disease, we have shown that PCR alone has sufficient DNA sequence specificity to permit genetic screening. Using appropriate amplification conditions, there is little non-specific production of dsDNA in the absence of the appropriate target allele.

The specificity required to detect pathogens can be more or less than that required to do genetic screening, depending on the number of pathogens in the sample and the amount of other DNA that must be taken with the sample. A difficult target is HIV, which requires detection of a viral genome that can be at the level of a few copies per thousands of host cells[9]. Compared with genetic screening, which is performed on cells containing at least one copy of the target sequence, HIV detection requires both more specificity and the input of more total

FIGURE 4 UV photography of PCR tubes containing amplifications using EtBr that are specific to wild-type (A) or sickle (S) alleles of the human β-globin gene. The left of each pair of tubes contains allele-specific primers to the wild-type allele, the right tube primers to the sickle allele. The photograph was taken after 30 cycles of PCR, and the input DNAs and the alleles they contain are indicated. Fifty ng of DNA was used to begin PCR. Typing was done in triplicate (3 pairs of PCRs) for each input DNA.

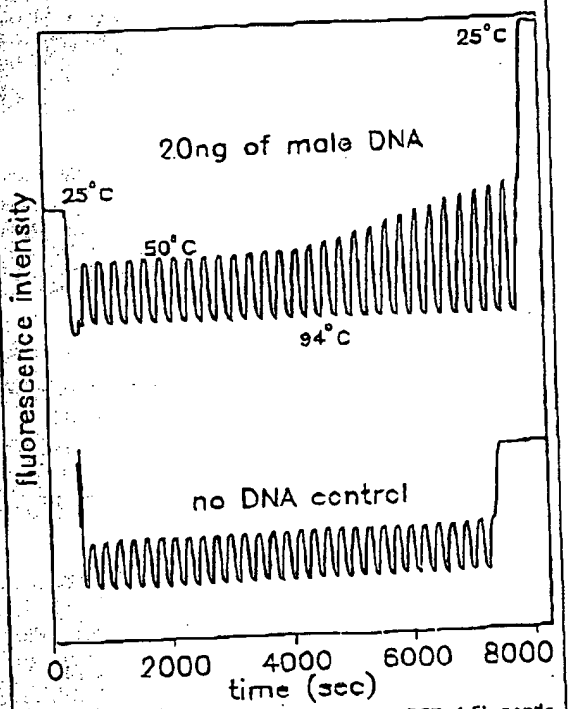


FIGURE 5 Continuous, real-time monitoring of a PCR. A fiber optic was used to carry excitation light to a PCR in progress and also emitted light back to a fluorometer (see Experimental Protocol). Amplification using human male-DNA specific primers in a PCR starting with 20 ng of human male DNA (top), or in a control PCR without DNA (bottom), were monitored. Thirty cycles of PCR were followed for each. The temperature cycled between 94°C (denaturation) and 50°C (annealing and extension). Note in the male DNA PCR, the cycle (time) dependent increase in fluorescence at the annealing/extension temperature.

DNA—up to microgram amounts—in order to have sufficient numbers of target sequences. This large amount of starting DNA in an amplification significantly increases the background fluorescence over which any additional fluorescence produced by PCR must be detected. An additional complication that occurs with targets in low copy-number is the formation of the "primer-dimer" artifact. This is the result of the extension of one primer using the other primer as a template. Although this occurs infrequently, once it occurs the extension product is a substrate for PCR amplification, and can compete with true PCR targets if those targets are rare. The primer-dimer product is of course dsDNA and thus is a potential source of false signal in this homogeneous assay.

To increase PCR specificity and reduce the effect of primer-dimer amplification, we are investigating a number of approaches, including the use of nested-primer amplifications that take place in a single tube[8], and the "hot-start", in which nonspecific amplification is reduced by raising the temperature of the reaction before DNA synthesis begins[22]. Preliminary results using these approaches suggest that primer-dimer is effectively reduced and it is possible to detect the increase in EtBr fluorescence in a PCR instigated by a single HIV genome in a background of 10^5 cells. With larger numbers of cells, the background fluorescence contributed by genomic DNA becomes problematic. To reduce this background, it may be possible to use sequence-specific DNA-binding dyes that can be made to preferentially bind PCR product over genomic DNA by incorporating the dye-binding DNA sequence into the PCR product through a 5' "add-on" to the oligonucleotide primer[24].

We have shown that the detection of fluorescence generated by an EtBr-containing PCR is straightforward, both once PCR is completed and continuously during thermocycling. The ease with which automation of specific DNA detection can be accomplished is the most promising aspect of this assay. The fluorescence analysis of completed PCRs is already possible with existing instrumentation in 96-well format[25]. In this format, the fluorescence in each PCR can be quantitated before, after, and even at selected points during thermocycling by moving the rack of PCRs to a 96-microwell plate fluorescence reader[26].

The instrumentation necessary to continuously monitor multiple PCRs simultaneously is also simple in principle. A direct extension of the apparatus used here is to have multiple fiberoptics transmit the excitation light and fluorescent emissions to and from multiple PCRs. The ability to monitor multiple PCRs continuously may allow quantitation of target DNA copy number. Figure 3 shows that the larger the amount of starting target DNA, the sooner during PCR a fluorescence increase is detected. Preliminary experiments (Higuchi and Dollinger, manuscript in preparation) with continuous monitoring have shown a sensitivity to two-fold differences in initial target DNA concentration.

Conversely, if the number of target molecules is known—as it can be in genetic screening—continuous monitoring may provide a means of detecting false positive and false negative results. With a known number of target molecules, a true positive would exhibit detectable fluorescence by a predictable number of cycles of PCR. Increases in fluorescence detected before or after that cycle would indicate potential artifacts. False negative results due to, for example, inhibition of DNA polymerase, may be detected by including within each PCR an inefficiently amplifying marker. This marker results in a fluorescence increase only after a large number of cycles—many more than are necessary to detect a true positive. If a sample fails to have a fluorescence increase after this many cycles, inhibition may be suspected. Since, in this assay, conclusions are drawn based on the presence or absence of fluorescence signal alone, such controls may be important. In any event, before any test based on this principle is ready for the clinic, an assessment of its false positive/false negative rates will need to be obtained using a large number of known samples.

In summary, the inclusion in PCR of dyes whose fluorescence is enhanced upon binding dsDNA makes it possible to detect specific DNA amplification from outside the PCR tube. In the future, instruments based upon this principle may facilitate the more widespread use of PCR in applications that demand the high throughput of samples.

EXPERIMENTAL PROTOCOL

Human HLA-DQα gene amplifications containing EtBr. PCRs were set up in 100 µl volumes containing 10 mM Tris-HCl, pH 8.3; 50 mM KCl; 4 mM $MgCl_2$; 2.5 units of *Taq* DNA polymerase (Perkin-Elmer Cetus, Norwalk, CT); 20 pmole each of human HLA-DQα gene specific oligonucleotide primers GH26 and GH27[19] and approximately 10^7 copies of DQα PCR product diluted from a previous reaction. Ethidium bromide (EtBr; Sigma) was used at the concentrations indicated in Figure 2. Thermocycling proceeded for 20 cycles in a model 480 thermocycler (Perkin-Elmer Cetus, Norwalk, CT) using a "step-cycle" program of 94°C for 1 min. denaturation and 60°C for 30 sec. annealing and 72°C for 30 sec. extension.

Y-chromosome specific PCR. PCRs (100 µl total reaction volume) containing 0.5 µg/ml EtBr were prepared as described for HLA-DQα, except with different primers and target DNAs. These PCRs contained 15 pmole each male DNA-specific primers Y1.1 and Y1.2[20], and either 60 ng male, 60 ng female, 2 ng male, or no human DNA. Thermocycling was 94°C for 1 min. and 60°C for 1 min using a "step-cycle" program. The number of cycles for a sample were as indicated in Figure 3. Fluorescence measurement is described below.

Allele-specific, human β-globin gene PCR. Amplifications of 100 µl volume using 0.5 µg/ml of EtBr were prepared as described for HLA-DQα above except with different primers and target DNAs. These PCRs contained either primer pair HGP2/Hβ14S (sickle-globin specific primers) or HGP2/Hβ14A (wild-type globin specific primers) at 10 pmole each primer per PCR. These primers were developed by Wu et al.[21]. Three different target DNAs were used in separate amplifications—50 ng each of human DNA that was homozygous for the sickle trait (SS), DNA that was heterozygous for the sickle trait (AS), or DNA that was homozygous for the w.t. globin (AA). Thermocycling was for 30 cycles at 94°C for 1 min. and 55°C for 1 min. using a "step-cycle" program. An annealing temperature of 55°C had been shown by Wu et al.[21] to provide allele-specific amplification. Completed PCRs were photographed through a red filter (Wratten 23A) after placing the reaction tubes upon a model TM-36 transilluminator (UV-products San Gabriel, CA).

Fluorescence measurement. Fluorescence measurements were made on PCRs containing EtBr in a Fluorolog-2 fluorometer (SPEX, Edison, NJ). Excitation was at the 500 nm band with about 2 nm bandwidth with a GG 485 nm cut-off filter (Melles Griot, Inc., Irvine, CA) to exclude second-order light. Emitted light was detected at 570 nm with a bandwidth of about 7 nm. An OG 530 nm cut-off filter was used to remove the excitation light.

Continuous fluorescence monitoring of PCR. Continuous monitoring of a PCR in progress was accomplished using the spectrofluorometer and settings described above as well as a fiberoptic accessory (SPEX cat. no. 1950) to both send excitation light to, and receive emitted light from, a PCR placed in a well of a model 480 thermocycler (Perkin-Elmer Cetus). The probe end of the fiberoptic cable was attached with "5 minute-epoxy" to the open top of a PCR tube (a 0.5 ml polypropylene centrifuge tube with its cap removed) effectively sealing it. The exposed top of the PCR tube and the end of the fiberoptic cable were shielded from room light and the room lights were kept dimmed during each run. The monitored PCR was an amplification of Y-chromosome-specific repeat sequences as described above, except using an annealing/extension temperature of 60°C. The reaction was covered with mineral oil (2 drops) to prevent evaporation. Thermocycling and fluorescence measurement were started simultaneously. A time-base scan with a 10 second integration time

was used and the emission signal was ratioed to the excitation signal to control for changes in light-source intensity. Data were collected using the dm3000f, version 2.6 (SPEX) data system.

Acknowledgments

We thank Bob Jones for help with the spectrofluormetric measurements and Heatherbell Fong for editing this manuscript.

References

1. Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. 1986. Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. CSLSQD 51:263–273.
2. White, T. J., Arnheim, N. and Erlich, H. A. 1989. The polymerase chain reaction. Trends Genet. 5:185–189.
3. Erlich, H. A., Gelfand, D. and Sninsky, J. J. 1991. Recent advances in the polymerase chain reaction. Science 252:1643–1651.
4. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491.
5. Saiki, R. K., Walsh, P. S., Levenson, C. H. and Erlich, H. A. 1989. Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. Proc. Nat. Acad. Sci. USA 86:6230–6234.
6. Kwok, S. Y., Mack, D. H., Mullis, K. B., Poiesz, B. J., Ehrlich, G. D., Blair, D. and Friedman-Kien, A. S. 1987. Identification of human immunodeficiency virus sequences by using *in vitro* enzymatic amplification and oligomer cleavage detection. J. Virol. 61:1690–1694.
7. Chehab, F. F., Doherty, M., Cai, S. P., Kan, Y. W., Cooper, S. and Rubin, E. M. 1987. Detection of sickle cell anemia and thalassemias. Nature 329:293–294.
8. Horn, G. T., Richards, B. and Klinger, K. W. 1989. Amplification of a highly polymorphic VNTR segment by the polymerase chain reaction. Nuc. Acids Res. 10:2140.
9. Katz, E. D. and Dong, M. W. 1990. Rapid analysis and purification of polymerase chain reaction products by high-performance liquid chromatography. Biotechniques 8:548–555.
10. Heiger, D. N., Cohen, A. S. and Karger, B. L. 1990. Separation of DNA restriction fragments by high performance capillary electrophoresis with low and zero crosslinked polyacrylamide using continuous and pulsed electric fields. J. Chromatogr. 516:33–48.
11. Kwok, S. Y. and Higuchi, R. G. 1989. Avoiding false positives with PCR. Nature 339:237–238.
12. Chehab, F. F. and Kan, Y. W. 1989. Detection of specific DNA sequences by fluorescence amplification: a color complementation assay. Proc. Nat. Acad. Sci. USA 86:9178–9182.
13. Holland, P. M., Abramson, R. D., Watson, R. and Gelfand, D. H. 1991. Detection of specific polymerase chain reaction product by utilizing the 5' to 3' exonuclease activity of *Thermus aquaticus* DNA polymerase. Proc. Nat. Acad. Sci. USA 88:7276–7280.
14. Markovic, J., Raques, B. P. and Le Pecq, J. B. 1979. Ethidium dimer: a new reagent for the fluorimetric determination of nucleic acids. Anal. Biochem. 94:259–264.
15. Kapuscinski, J. and Szer, W. 1979. Interactions of 4',6-diamidino-2-phenylindole with synthetic polynucleotides. Nuc. Acids Res. 6:3519–3534.
16. Searle, M. S. and Embrey, K. J. 1990. Sequence-specific interaction of Hoechst 33258 with the minor groove of an adenine-tract DNA duplex studied in solution by ^{1}H NMR spectroscopy. Nuc. Acids Res. 18:3753–3762.
17. Li, H. H., Gyllensten, U. B., Cui, X. F., Saiki, R. K., Erlich, H. A. and Arnheim, N. 1988. Amplification and analysis of DNA sequences in single human sperm and diploid cells. Nature 335:414–417.
18. Abbott, M. A., Poiesz, B. J., Byrne, B. C., Kwok, S. Y., Sninsky, J. J. and Erlich, H. A. 1988. Enzymatic gene amplification: qualitative and quantitative methods for detecting proviral DNA amplified *in vitro*. J. Infect. Dis. 158:1158.
19. Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B. and Erlich, H. A. 1986. Analysis of enzymatically amplified β-globin and HLA-DQα DNA with allele-specific oligonucleotide probes. Nature 324:163–166.
20. Kogan, S. C., Doherty, M. and Gitschier, J. 1987. An improved method for prenatal diagnosis of genetic diseases by analysis of amplified DNA sequences. N. Engl. J. Med. 317:985–990.
21. Wu, D. Y., Ugozzoli, L., Pal, B. K. and Wallace, R. B. 1989. Allele-specific enzymatic amplification of β-globin genomic DNA for diagnosis of sickle cell anemia. Proc. Nat. Acad. Sci. USA 86:2757–2760.
22. Kwok, S., Kellogg, D. E., McKinney, N., Spasic, D., Goda, L., Levenson, C. and Sninsky, J. J. 1990. Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies. Nuc. Acids Res. 18:999–1005.
23. Chou, Q., Russell, M., Birch, D., Raymond, J. and Bloch, W. 1992. Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. Submitted.
24. Higuchi, R. 1989. Using PCR to engineer DNA, p. 61–70. *In*: PCR Technology. H. A. Erlich (Ed.). Stockton Press. New York, N.Y.
25. Huff, L., Atwood, J. G., DiCesare, J., Katz, E., Fiscus, E., Williams, J. P. and Woudenberg, T. 1991. A high-performance system for automation of the polymerase chain reaction. Biotechniques 10:102–108, 110–112.
26. Tumosa, N. and Kahan, L. 1989. Fluorescent EIA screening of monoclonal antibodies to cell surface antigens. J. Immun. Meth. 116:59–68.

HellerEhrman
ATTORNEYS

Research

Oligonucleotides with Fluorescent Dyes at Opposite Ends Provide a Quenched Probe System Useful for Detecting PCR Product and Nucleic Acid Hybridization

Kenneth J. Livak, Susan J.A. Flood, Jeffrey Marmaro, William Giusti, and Karin Deetz

Perkin-Elmer, Applied Biosystems Division, Foster City, California 94404

The 5′ nuclease PCR assay detects the accumulation of specific PCR product by hybridization and cleavage of a double-labeled fluorogenic probe during the amplification reaction. The probe is an oligonucleotide with both a reporter fluorescent dye and a quencher dye attached. An increase in reporter fluorescence intensity indicates that the probe has hybridized to the target PCR product and has been cleaved by the 5′→3′ nucleolytic activity of *Taq* DNA polymerase. In this study, probes with the quencher dye attached to an internal nucleotide were compared with probes with the quencher dye attached to the 3′-end nucleotide. In all cases, the reporter dye was attached to the 5′ end. All intact probes showed quenching of the reporter fluorescence. In general, probes with the quencher dye attached to the 3′-end nucleotide exhibited a larger signal in the 5′ nuclease PCR assay than the internally labeled probes. It is proposed that the larger signal is caused by increased likelihood of cleavage by *Taq* DNA polymerase when the probe is hybridized to a template strand during PCR. Probes with the quencher dye attached to the 3′-end nucleotide also exhibited an increase in reporter fluorescence intensity when hybridized to a complementary strand. Thus, oligonucleotides with reporter and quencher dyes attached at opposite ends can be used as homogeneous hybridization assays.

A homogeneous assay for detecting the accumulation of specific PCR product that uses a double-labeled fluorogenic probe was described by Lee et al.[1] The assay exploits the 5′→3′ nucleolytic activity of *Taq* DNA polymerase[2,3] and is diagramed in figure 1. The fluorogenic probe consists of an oligonucleotide with a reporter fluorescent dye, such as a fluorescein, attached to the 5′ end, and a quencher dye, such as a rhodamine, attached internally. When the fluorescein is excited by irradiation, its fluorescent emission will be quenched if the rhodamine is close enough to be excited through the process of fluorescence energy transfer (FET).[4,5] During PCR, if the probe is hybridized to a template strand, *Taq* DNA polymerase will cleave the probe because of its inherent 5′→3′ nucleolytic activity. If the cleavage occurs between the fluorescein and rhodamine dyes, it causes an increase in fluorescein fluorescence intensity because the fluorescein is no longer quenched. The increase in fluorescein fluorescence intensity indicates that the probe-specific PCR product has been generated. Thus, FRT between a reporter dye and a quencher dye is critical to the performance of the probe in the 5′ nuclease PCR assay.

Quenching is completely dependent on the physical proximity of the two dyes.[4,5] Because of this, it has been assumed that the quencher dye must be attached near the 5′ end. Surprisingly, we have found that attaching a rhodamine dye at the 3′ end of a probe for the 5′ nuclease PCR assay. Furthermore, cleavage of this type of probe is not required to achieve some reduction in quenching. Oligonucleotides with a reporter dye on the 5′ end and a quencher dye on the 3′ end exhibit a much higher reporter fluorescence when double-stranded as compared with single-stranded. This should make it possible to use this type of double-labeled probe for homogeneous detection of nucleic acid hybridization.

MATERIALS AND METHODS

Oligonucleotides

Table 1 shows the nucleotide sequence of the oligonucleotides used in this study. Linker arm nucleotide (LAN) phosphoramidite was obtained from Glen Research. The standard DNA phosphoramidites, 6-carboxyfluorescein (6-FAM) phosphoramidite, 6-carboxytetramethylrhodamine succinimidyl ester (TAMRA NHS ester), and Phosphalink for attaching a 3′-blocking phosphate, were obtained from Perkin-Elmer, Applied Biosystems Division. Oligonucleotide synthesis was performed using an ABI model 394 DNA synthesizer (Applied Biosystems). Primer and complement oligonucleotides were purified using Oligo Purification Cartridges (Applied Biosystems). Double-labeled probes were synthesized with 6-FAM-labeled phosphoramidite at the 5′ end, LAN replacing one of the T's in the sequence, and Phosphalink at the 3′ end. Following deprotection and ethanol precipitation,

Research

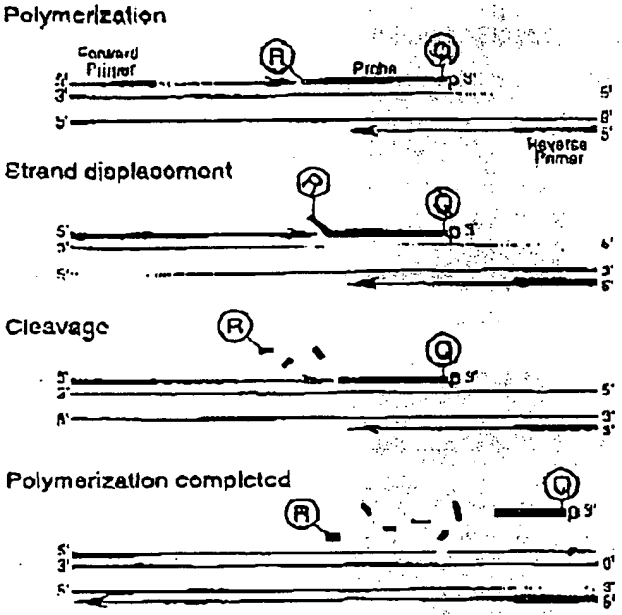


FIGURE 1 Diagram of 5′ nuclease assay. Stepwise representation of the 5′ → 3′ nucleolytic activity of *Taq* DNA polymerase acting on a fluorogenic probe during one extension phase of PCR.

mM Na-bicarbonate buffer (pH 9.0) at room temperature. Unreacted dye was removed by passage over a PD-10 Sephadex column. Finally, the double-labeled probe was purified by preparative high-performance liquid chromatography (HPLC) using an Aquapore C_{18} 220×4.6-mm column with 7-μm particle size. The column was developed with a 24-min linear gradient of 8–20% acetonitrile in 0.1 M TEAA (triethylamine acetate). Probes are named by designating the sequence from Table 1 and the position of the TAN–TAMRA moiety. For example, probe A1-7 has sequence A1 with TAN–TAMRA at nucleotide position 7 from the 5′ end.

PCR Systems

All PCR amplifications were performed in the Perkin-Elmer GeneAmp PCR System 9600 using 50-μl reactions that contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 200 μM dATP, 200 μM dCTP, 200 μM dGTP, 400 μM dUTP, 0.5 unit of AmpErase uracil *N*-glycosylase (Perkin-Elmer), gene (nucleotides 2141–2435 in the sequence of Nakajima-Iijima et al.)[3] was amplified using primers AFP and ARP (Table 1), which are modified slightly from those of du Breuil et al.[4] Actin amplification reactions contained 4 mM $MgCl_2$, 20 ng of human genomic DNA, 50 nM A1 or A2 probe, and 300 nM each primer. The thermal regimen was 50°C (2 min), 95°C (10 min), 40 cycles of 95°C (20 sec), 60°C (1 min), and hold at 72°C. A 515-bp segment was amplified from a plasmid that consists of a segment of λ DNA (nucleotides 32,220–32,747) inserted in the *Sma*I site of vector pUC119. These reactions contained 5.5 mM $MgCl_2$, 1 ng of plasmid DNA, 50 nM P2 or P5 probe, 200 nM primer P119, and 300 nM primer R119. The thermal regimen was 50°C (2 min), 95°C (10 min), 25 cycles of 95°C (20 sec), 57°C (1 min), and hold at 72°C.

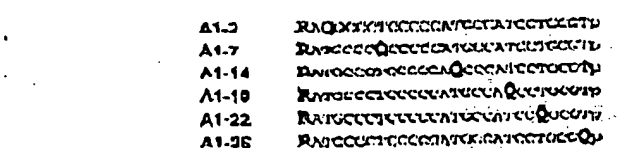
Fluorescence Detection

For each amplification reaction, a 40-μl aliquot of a sample was transferred to an individual well of a white, 96-well microtiter plate (Perkin-Elmer). Fluorescence was measured on the Perkin-Elmer TaqMan LS-50B System, which consists of a luminescence spectrometer with plate reader assembly, a 488-nm excitation filter, and a 515-nm emission filter. Excitation was at 488 nm using a 5-nm slit width. Emission was measured at 518 nm for 6-FAM (the reporter or R value) and 582 nm for TAMRA (the quencher or Q value) using a 10-nm slit width. To determine the increase in reporter emission that is caused by cleavage of the probe during PCR, three normalizations are applied to the raw emission data. First, emission intensity of a buffer blank is subtracted for each wavelength. Second, emission intensity of the reporter is

TABLE 1 Sequences of Oligonucleotides

Name	Type	Sequence
P119	primer	ACCCACAGGAACTGATCACCACTC
R119	primer	ATGTCGCGTTTCCCGCTCACCATTCTTC
P2	probe	TCCCATTACTGATCGTTTCCCAACCACTp
P2C	complement	GTACTCGTTGGGAACCATCAGTAATGCCATG
P5	probe	CCCATTTGCTCGTATCTATCACAACCATp
P5C	complement	TTCATCGTTCTCATAGATACCACAAATCCC
AFP	primer	TCACCCACACTGTGCCCATCTACGA
ARP	primer	CAGCGGAACCGTCTCATTCCCAATCC
A1	probe	ATCCCCCCCCCCATGCCATCCTCCCGTp
A1C	complement	ACACCCACCATCCCATCCCCCCACCCCCATAC
A2	probe	CGCCCTGGACTTCCACCAACACACATp
A2C	complement	CCATCTCTTCCTCCAACTCCACCCCCAAC

For each oligonucleotide used in this study, the nucleic acid sequence is given, written in the 5′ → 3′ direction. There are three types of oligonucleotides: PCR primer, fluorogenic probe used



Probe	518 nm		582 nm		RQ⁻	RQ⁺	ΔRQ
	no temp.	+ temp.	no temp.	+ temp.			
A1-2	38.6 ± 2.1	32.7 ± 1.0	66.2 ± 0.0	68.0 ± 2.0	0.47 ± 0.01	0.30 ± 0.02	0.16 ± 0.02
A1-7	83.8 ± 4.3	306.1 ± 21.4	108.5 ± 6.0	110.3 ± 6.3	0.40 ± 0.02	0.58 ± 0.17	2.00 ± 0.18
A1-14	127.0 ± 4.9	433.5 ± 16.1	108.7 ± 5.3	93.1 ± 6.3	1.18 ± 0.03	4.24 ± 0.18	3.18 ± 0.15
A1-19	197.5 ± 19.9	499.7 ± 7.7	70.3 ± 7.4	78.0 ± 0.0	2.67 ± 0.06	5.00 ± 0.16	2.12 ± 0.16
A1-22	224.0 ± 0.4	480.0 ± 43.6	100.0 ± 4.0	96.2 ± 0.6	2.25 ± 0.03	5.02 ± 0.11	2.77 ± 0.12
A1-28	160.2 ± 0.9	454.1 ± 18.4	93.1 ± 5.4	90.7 ± 3.2	1.72 ± 0.02	5.01 ± 0.05	3.29 ± 0.05

FIGURE 2 Results of 5′ nuclease assay comparing β-actin probes with TAMRA at different nucleotide positions. As described in Materials and Methods, PCR amplifications containing the indicated probes were performed, and the fluorescence emission was measured at 518 and 582 nm. Reported values are the average±1 s.d. for six reactions run without added template (no temp.) and six reactions run with template (+ temp.). The RQ ratio was calculated for each individual reaction and averaged to give the reported RQ⁻ and RQ⁺ values.

divided by the emission intensity of the quencher to give an RQ ratio for each reaction tube. This normalizes for well-to-well variations in probe concentration and fluorescence measurement. Finally, ΔRQ is calculated by subtracting the RQ value of the no-template control (RQ^-) from the RQ value for the complete reaction including template (RQ^+).

RESULTS

A series of probes with increasing distances between the fluorescent reporter and rhodamine quencher were tested to investigate the minimum and maximum spacing that would give an acceptable performance in the 5′ nuclease PCR assay. These probes hybridize to a target sequence in the human β-actin gene. Figure 2 shows the results of an experiment in which these probes were included in PCR that amplified a segment of the β-actin gene containing the target sequence. Performance in the 5′ nuclease PCR assay is monitored by the magnitude of ΔRQ, which is a measure of the increase in reporter fluorescence caused by PCR amplification of the probe target. Probe A1-2 has a ΔRQ value that is close to zero, indicating that the probe was not cleaved appreciably during the amplification reaction. This suggests that with the quencher dye on the second nucleotide from the 5′ end, there is insufficient room for *Taq* polymerase to cleave efficiently between the reporter and quencher. The other five probes exhibited comparable ΔRQ values that are clearly different from zero. Thus, all five probes are being cleaved during PCR amplification resulting in a similar increase in reporter fluorescence. It should be noted that complete digestion of a probe produces a much larger increase in reporter fluorescence than that observed in Figure 2 (data not shown). Thus, even in reactions where amplification occurs, the majority of probe molecules remain uncleaved. It is mainly for this reason that the fluorescence intensity of the quencher dye TAMRA changes little with amplification of the target. This is what allows us to use the 582-nm fluorescence reading as a normalization factor.

The magnitude of RQ^- depends mainly on the quenching efficiency inherent in the specific structure of the probe and the purity of the oligonucleotide. Thus, the larger RQ^- values indicate that probes A1-14, A1-19, A1-22, and A1-28 probably have reduced quenching as compared with A1-7. Still, the degree of quenching is sufficient to detect a highly significant increase in reporter fluorescence when each of these probes is cleaved during PCR.

To further investigate the ability of TAMRA on the 3′ end to quench 6-FAM on the 5′ end, three additional pairs of probes were tested in the 5′ nuclease PCR assay. For each pair, one probe has TAMRA attached to an internal nucleotide and the other has TAMRA attached to the 3′ end nucleotide. The results are shown in Table 2. For all three sets, the probe with the 3′ quencher exhibits a ΔRQ value that is considerably higher than for the probe with the internal quencher. The RQ^+ values suggest that differences in quenching are not as great as those observed with some of the A1 probes. These results demonstrate that a quencher dye on the 3′ end of an oligonucleotide can quench efficiently the

TABLE 2 Results of 5′ Nuclease Assay Comparing Probes with TAMRA Attached to an Internal or 3′-terminal Nucleotide

Probe	518 nm		582 nm		RQ	RQ⁺	ΔRQ
	no temp.	+ temp.	no temp.	+ temp.			
A3-6	54.6 ± 3.2	84.8 ± 3.7	116.2 ± 6.4	113.6 ± 2.5	0.47 ± 0.02	0.73 ± 0.03	0.26 ± 0.04
A3-24	72.1 ± 2.9	236.5 ± 11.1	84.2 ± 4.0	90.2 ± 3.8	0.86 ± 0.02	2.62 ± 0.05	1.76 ± 0.05
P2-7	82.8 ± 4.4	384.0 ± 34.1	105.1 ± 6.4	120.4 ± 10.2	0.79 ± 0.02	3.19 ± 0.16	2.40 ± 0.16
P2-27	113.4 ± 6.6	556.4 ± 14.1	140.7 ± 8.3	118.7 ± 4.8	0.81 ± 0.01	4.68 ± 0.10	3.68 ± 0.10
P3-10	77.3 ± 6.5	244.4 ± 15.9	86.7 ± 4.4	95.8 ± 6.7	0.89 ± 0.08	2.55 ± 0.06	1.66 ± 0.08
P3-28	64.0 ± 3.2	333.6 ± 12.1	100.6 ± 8.1	94.7 ± 6.3	0.63 ± 0.02	3.54 ± 0.12	2.89 ± 0.13

* —Reactions were performed as described in Material and Methods and in the legend to Fig. 2.

Research

fluorescence of a reporter dye on the 5′ end. The degree of quenching is sufficient for this type of oligonucleotide to be used as a probe in the 5′ nuclease PCR assay.

To test the hypothesis that quenching by a 3′ TAMRA depends on the flexibility of the oligonucleotide, fluorescence was measured for probes in the single-stranded and double-stranded states. Table 3 reports the fluorescence observed at 518 and 582 nm. The relative degree of quenching is assessed by calculating the RQ ratio. For probes with TAMRA 6–10 nucleotides from the 5′ end, there is little difference in the RQ values when comparing single-stranded with double-stranded oligonucleotides. The results for probes with TAMRA at the 3′ end are much different. For these probes, hybridization to a complementary strand causes a dramatic increase in RQ. We propose that this loss of quenching is caused by the rigid structure of double-stranded DNA, which prevents the 5′ and 3′ ends from being in proximity.

When TAMRA is placed toward the 3′ end, there is a marked Mg^{2+} effect on quenching. Figure 3 shows a plot of observed RQ values for the A1 series of probes as a function of Mg^{2+} concentration. With TAMRA attached near the 5′ end (probe A1-2 or A1-7), the RQ value at 0 mM Mg^{2+} is only slightly higher than RQ at 10 mM Mg^{2+}. For probes A1-19, A1-22, and A1-26, the RQ values at 0 mM Mg^{2+} are very high, indicating a much reduced quenching efficiency. For each of these probes, there is a marked decrease in RQ at 3 mM Mg^{2+} followed by a gradual decline as the Mg^{2+} concentration increases to 10 mM. Probe A1-14 shows an intermediate RQ value at 0 mM Mg^{2+} with a gradual decline at higher Mg^{2+} concentrations. In a low-salt environment with no Mg^{2+} present, a single-stranded oligonucleotide would be expected to adopt an extended conformation because of electrostatic repulsion. The binding of Mg^{2+} ions acts to shield the negative charge of the phosphate backbone so that the oligonucleotide can adopt conformations where the 3′ end is close to the 5′ end. Therefore, the observed Mg^{2+} effects support the notion that quenching of a 5′ reporter dye by TAMRA at or near the 3′ end depends on the flexibility of the oligonucleotide.

DISCUSSION

The striking finding of this study is that it seems the rhodamine dye TAMRA, placed at any position in an oligonucleotide, can quench the fluorescent emission of a fluorescein (6-FAM) placed at the 5′ end. This implies that a single-stranded, double-labeled oligonucleotide must be able to adopt conformations where the TAMRA is close to the 5′ end. It should be noted that the decay of 6-FAM in the excited state requires a certain amount of time. Therefore, what matters for quenching is not the average distance between 6-FAM and TAMRA but, rather, how close TAMRA can get to 6-FAM during the lifetime of the 6-FAM excited state. As long as the decay time of the excited state is relatively long compared with the molecular motions of the oligonucleotide, quenching can occur. Thus, we propose that TAMRA at the 3′ end, or any other position, can quench 6-FAM at the 5′ end because TAMRA is in proximity to 6-FAM often enough to be able to accept energy transfer from an excited 6-FAM.

Details of the fluorescence measurements remain puzzling. For example, Table 3 shows that hybridization of probes A1-26, A3-24, and P5-28 to their complementary strands not only causes a large increase in 6-FAM fluorescence at 518 nm but also causes a modest increase in TAMRA fluorescence at 582 nm. If TAMRA is being excited by energy transfer from quenched 6-FAM, then loss of quenching attributable to hybridization should cause a decrease in the fluorescence emission of TAMRA. The fact that the fluorescence emission of TAMRA increases indicates that the situation is more complex. For example, we have anecdotal evidence that the bases of the oligonucleotide, especially G, quench the fluorescence of both 6-FAM and TAMRA to some degree. When double-stranded, base-pairing may reduce the ability of the bases to quench. The primary factor causing the quenching of 6-FAM in an intact probe is the TAMRA dye. Evidence for the importance of TAMRA is that 6-FAM fluorescence remains relatively unchanged when probes labeled only with 6-FAM are used in the 5′ nuclease PCR assay (data not shown). Secondary effectors of fluorescence, both before and after cleavage of the probe, need to be explored further.

Regardless of the physical mechanism, the relative independence of position and quenching greatly simplifies the design of probes for the 5′ nuclease PCR assay. There are three main factors that determine the performance of a double-labeled fluorescent probe in the 5′ nuclease PCR assay. The first factor is the degree of quenching observed in the intact probe. This is characterized by the value of RQ′, which is the ratio of reporter to quencher fluorescent emis-

TABLE 3 Comparison of Fluorescence Emissions of Single-stranded and Double-stranded Fluorogenic Probes

	518 nm		582 nm		RQ	
Probe	ss	ds	ss	ds	ss	ds
A1-7	27.75	68.59	61.08	138.18	0.45	11.50
A1-26	43.51	309.38	53.50	95.86	0.81	5.43
A3-6	16.75	67.88	39.11	165.57	0.43	0.38
A3-24	30.05	578.64	67.72	140.25	0.45	3.21
P2-7	35.02	70.13	54.63	121.09	0.64	0.58
P2-27	20.80	220.47	65.10	61.13	0.61	5.25
P5-10	27.14	144.85	61.95	165.54	0.44	0.87
P5-28	33.68	462.20	92.30	104.41	0.46	4.43

(ss) Single-stranded. The fluorescence emissions at 518 or 582 nm for solutions containing a final concentration of 50 nM indicated probe, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, and 10 mM $MgCl_2$. (ds) Double-stranded. The solutions contained, in addition, 100 nM A1C for probes A1-7 and A1-26, 100 nM A3C for probes A3-6 and A3-24, 100 nM P2C for probes P2-7 and P2-27, or 100 nM P5C for probes P5-10 and P5-28. Before the addition of $MgCl_2$, 120 μl of each sample was heated

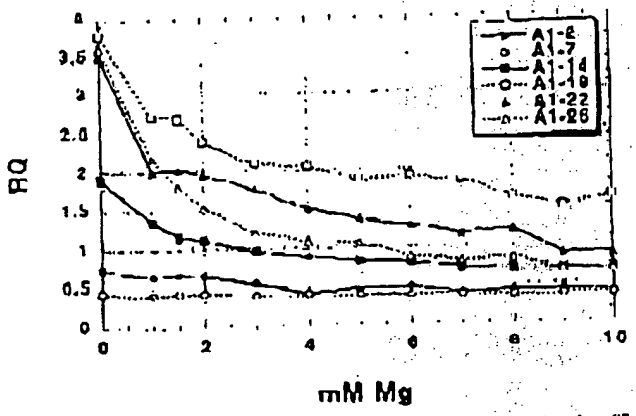


FIGURE 3 Effect of Mg^{++} concentration on RQ ratio for the A1 series of probes. The fluorescence emission intensity at 518 and 582 nm was measured for solutions containing 50 nM probe, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, and varying amounts (0–10 mM) of $MgCl_2$. The calculated RQ ratios (518 nm intensity divided by 582 nm intensity) are plotted vs. $MgCl_2$ concentration (mM Mg). The key (*upper right*) shows the probes examined.

dyes used, spacing between reporter and quencher dyes, nucleotide sequence context effects, presence of structure or other factors that reduce flexibility of the oligonucleotide, and purity of the probe. The second factor is the efficiency of hybridization, which depends on probe T_m, presence of secondary structure in probe or template, annealing temperature, and other reaction conditions. The third factor is the efficiency at which *Taq* DNA polymerase cleaves the bound probe between the reporter and quencher dyes. This cleavage is dependent on sequence complementarity between probe and template as shown by the observation that mismatches in the segment between reporter and quencher dyes drastically reduce the cleavage of probe.[4]

The rise in RQ$^-$ values for the A1 series of probes seems to indicate that the degree of quenching is reduced somewhat as the quencher is placed toward the 3′ end. The lowest apparent quenching is observed for probe A1-19 (see Fig. 3) rather than for the probe where the TAMRA is at the 3′ end (A1-26). This is understandable, as the conformation of the 3′ end position would be expected to be less restricted than the conformation of an internal position. In effect, a quencher at the 3′ end is freer to adopt conformations close to the 5′ reporter dye than is an internally placed probes, the interpretation of RQ$^-$ values is less clear-cut. The A3 probes show the same trend as A1, with the 3′ TAMRA probe having a larger RQ$^-$ than the internal TAMRA probe. For the P2 pair, both probes have about the same RQ$^-$ value. For the P5 probes, the RQ$^-$ for the 3′ probe is less than for the internally labeled probe. Another factor that may explain some of the observed variation is that purity affects the RQ$^-$ value. Although all probes are HPLC purified, a small amount of contamination with unquenched reporter can have a large effect on RQ$^-$.

Although there may be a modest effect on degree of quenching, the position of the quencher apparently can have a large effect on the efficiency of probe cleavage. The most drastic effect is observed with probe A1-2, where placement of the TAMRA on the second nucleotide reduces the efficiency of cleavage to almost zero. For the A3, P2, and P5 probes, ΔRQ is much greater for the 3′ TAMRA probes as compared with the internal TAMRA probes. This is explained most easily by assuming that probes with TAMRA at the 3′ end are more likely to be cleaved between reporter and quencher than are probes with TAMRA attached internally. For the A1 probes, the cleavage efficiency of probe A1-7 must already be quite high, as ΔRQ does not increase when the quencher is placed closer to the 3′ end. This illustrates the importance of being able to use probes with a quencher on the 3′ end in the 5′ nuclease PCR assay. In this assay, an increase in the intensity of reporter fluorescence is observed only when the probe is cleaved between the reporter and quencher dyes. By placing the reporter and quencher dyes on the opposite ends of an oligonucleotide probe, any cleavage that occurs will be detected. When the quencher is attached to an internal nucleotide, sometimes the probe works well (A1-7) and other times not so well (A3-6). The relatively poor performance of probe A3-6 presumably means the probe is being cleaved 3′ to the quencher rather than between the reporter and quencher. Therefore, the best chance of having a probe that reliably detects accumulation of PCR product in the 5′ nuclease PCR assay is to use a probe with the reporter and quencher dyes on opposite ends.

Placing the quencher dye on the 3′ end may also provide a slight benefit in terms of hybridization efficiency. The presence of a quencher attached to an internal nucleotide might be expected to disrupt base-pairing and reduce the T_m of a probe. In fact, a 2°C–3°C reduction in T_m has been observed for two probes with internally attached TAMRAs.[10] This disruptive effect would be minimized by placing the quencher at the 3′ end. Thus, probes with 3′ quenchers might exhibit slightly higher hybridization efficiencies than probes with internal quenchers.

The combination of increased cleavage and hybridization efficiencies means that probes with 3′ quenchers probably will be more tolerant of mismatches between probe and target as compared with internally labeled probes. This tolerance of mismatches can be advantageous, as when trying to use a single probe to detect PCR-amplified products from samples of different species. Also, it means that cleavage of probe during PCR is less sensitive to alterations in annealing temperature or other reaction conditions. The one application where tolerance of mismatches may be a disadvantage is for allelic discrimination. Lee et al.[3] demonstrated that allele-specific probes were cleaved between reporter and quencher only when hybridized to a perfectly complementary target. This allowed them to distinguish the normal human cystic fibrosis allele from the *ΔF508* mutant. Their probes had TAMRA attached to the seventh nucleotide from

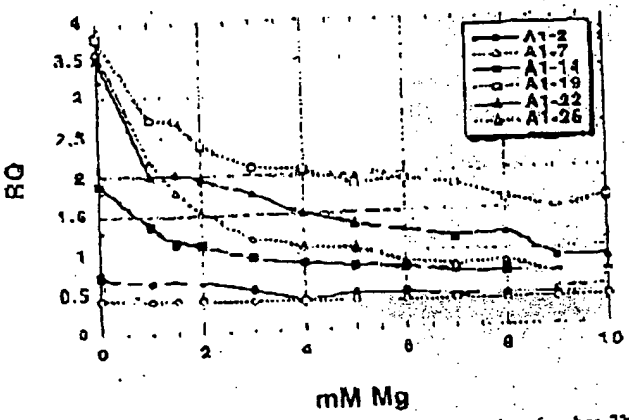


FIGURE 3 Effect of Mg^{2+} concentration on RQ ratio for the A1 series of probes. The fluorescence emission intensity at 518 and 582 nm was measured for solutions containing 50 nM probe, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, and varying amounts (0–10 mM) of $MgCl_2$. The calculated RQ ratios (518 nm intensity divided by 582 nm intensity) are plotted vs. $MgCl_2$ concentration (mM Mg). The key (*upper right*) shows the probes examined.

dyes used, spacing between reporter and quencher dyes, nucleotide sequence context effects, presence of structure or other factors that reduce flexibility of the oligonucleotide, and purity of the probe. The second factor is the efficiency of hybridization, which depends on probe T_m, presence of secondary structure in probe or template, annealing temperature, and other reaction conditions. The third factor is the efficiency at which *Taq* DNA polymerase cleaves the bound probe between the reporter and quencher dyes. This cleavage is dependent on sequence complementarity between probe and template as shown by the observation that mismatches in the segment between reporter and quencher dyes drastically reduce the cleavage of probe.[5]

The rise in RQ values for the A1 series of probes seems to indicate that the degree of quenching is reduced somewhat as the quencher is placed toward the 3' end. The lowest apparent quenching is observed for probe A1-19 (see Fig. 3) rather than for the probe where the TAMRA is at the 3' end (A1-26). This is understandable, as the conformation of the 3' end position would be expected to be less restricted than the conformation of an internal position. In effect, a quencher at the 3' end is freer to adopt conformations close to the 5' reporter dye than is an internally placed [illegible] For the other three sets of probes, the interpretation of RQ values is less clear-cut. The A3 probes show the same trend as A1, with the 3' TAMRA probe having a larger RQ than the internal TAMRA probe. For the P2 pair, both probes have about the same RQ value. For the P5 probes, the RQ for the 3' probe is less than for the internally labeled probe. Another factor that may explain some of the observed variation is that partly affects the RQ value. Although all probes are HPLC purified, a small amount of contamination with unquenched reporter can have a large effect on RQ.

Although there may be a modest effect on degree of quenching, the position of the quencher apparently can have a large effect on the efficiency of probe cleavage. The most drastic effect is observed with probe A1-2, where placement of the TAMRA on the second nucleotide reduces the efficiency of cleavage to almost zero. For the A3, P2, and P5 probes, ΔRQ is much greater for the 3' TAMRA probes as compared with the internal TAMRA probes. This is explained most easily by assuming that probes with TAMRA at the 3' end are more likely to be cleaved between reporter and quencher than are probes with TAMRA attached internally. For the A1 probes, the cleavage efficiency of probe A1-7 must already be quite high, as ΔRQ does not increase when the quencher is placed closer to the 3' end. This illustrates the importance of being able to use probes with a quencher on the 3' end in the 5' nuclease PCR assay. In this assay, an increase in the intensity of reporter fluorescence is observed only when the probe is cleaved between the reporter and quencher dyes. By placing the reporter and quencher dyes on the opposite ends of an oligonucleotide probe, any cleavage that occurs will be detected. When the quencher is attached to an internal nucleotide, sometimes the probe works well (A1-7) and other times not so well (A3-6). The relatively poor performance of probe A2-6 presumably means the probe is being cleaved 3' to the quencher rather than between the reporter and quencher. Therefore, the best chance of having a probe that reliably detects accumulation of PCR product in the 5' nuclease PCR assay is to use a probe with the reporter and quencher dyes on opposite ends.

Placing the quencher dye on the 3' end may also provide a slight benefit in terms of hybridization efficiency. The presence of a quencher attached to an internal nucleotide might be expected to disrupt base-pairing and reduce the T_m of a probe. In fact, a 2°C–3°C reduction in T_m has been observed for two probes with internally attached TAMRAs.[9] This disruptive effect would be minimized by placing the quencher at the 3' end. Thus, probes with 3' quenchers might exhibit slightly higher hybridization efficiencies than probes with internal quenchers.

The combination of increased cleavage and hybridization efficiencies means that probes with 3' quenchers probably will be more tolerant of mismatches between probe and target as compared with internally labeled probes. This tolerance of mismatches can be advantageous, as when trying to use a single probe to detect PCR-amplified products from samples of different species. Also, it means that cleavage of probe during PCR is less sensitive to alterations in annealing temperature or other reaction conditions. One application where tolerance of mismatches may be a disadvantage is for allelic discrimination. Lee et al.[10] demonstrated that allele-specific probes were cleaved between reporter and quencher only when hybridized to a perfectly complementary target. This allowed them to distinguish the normal human cystic fibrosis allele from the *ΔF508* mutant. Their probes had TAMRA attached to the seventh nucleotide from

the 5' end and were designed so that any mismatches were between the reporter and quencher. Increasing the distance between reporter and quencher would lessen the disruptive effect of mismatches and allow cleavage of the probe on the incorrect target. Thus, probes with a quencher attached to an internal nucleotide may still be useful for allelic discrimination.

In this study loss of quenching upon hybridization was used to show that quenching by a 3' TAMRA is dependent on the flexibility of a single-stranded oligonucleotide. The increase in reporter fluorescence intensity, though, could also be used to determine whether hybridization has occurred or not. Thus, oligonucleotides with reporter and quencher dyes attached at opposite ends should also be useful as hybridization probes. The ability to detect hybridization in real time means that these probes could be used to measure hybridization kinetics. Also, this type of probe could be used to develop homogeneous hybridization assays for diagnostics or other applications. Bagwell et al.[10] describe just this type of homogeneous assay where hybridization of a probe causes an increase in fluorescence caused by a loss of quenching. However, they utilized a complex probe design that requires adding nucleotides to both ends of the probe sequence to form two imperfect hairpins. The results presented here demonstrate that the simple addition of a reporter dye to one end of an oligonucleotide and a quencher dye to the other end generates a fluorogenic probe that can detect hybridization or PCR amplification.

ACKNOWLEDGMENTS

We acknowledge Lincoln McBride of Perkin-Elmer for his support and encouragement on this project and Mitch Winnik of the University of Toronto for helpful discussions on time-resolved fluorescence.

REFERENCES

1. Lee, L.G., C.R. Connell, and W. Bloch. 1993. Allelic discrimination by nick-translation PCR with fluorogenic probes. *Nucleic Acids Res.* 21: 3761–3766.
2. [illegible] 1991. [illegible]
3. Lyamichev, V., M.A.D. Brow, and J.E. Dahlberg. 1993. Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerases. *Science* 260: 778–783.
4. Förster, V.Th. 1948. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys. (Leipzig)* 2: 55–75.
5. Lakowicz, J.R. 1983. Energy transfer. In *Principles of fluorescent spectroscopy.* pp. 303–339. Plenum Press, New York, NY.
6. Stryer, L. and R.P. Haugland. 1967. Energy transfer: A spectroscopic ruler. *Proc. Natl. Acad. Sci. USA* 58: 719–726.
7. Nakajima-Iijima, S., H. Hamada, P. Reddy, and T. Kakunaga. 1985. Molecular structure of the human cytoplasmic beta-actin gene: Interspecies homology of sequences in the introns. *Proc. Natl. Acad. Sci.* 82: 6133–6137.
8. du Breuil, R.M., J.M. Patel, and R.V. Meudelow. 1993. Quantitation of β-actin-specific mRNA transcripts using semi-competitive PCR. *PCR Methods Applic.* 3: 57–59.
9. Livak, K.J. (unpubl.).
10. Bagwell, C.B., M.E. Munson, R.L. Christensen, and B.J. Lovett. 1994. A new homogeneous assay system for specific nucleic acid sequences, Poly-dA and poly-A detection. *Nucleic Acids Res.* 22: 2424–2425.

Received December 30, 1994; accepted in revised form March 6, 1995.

HellerEhrman
ATTORNEYS

RESEARCH

SIMULTANEOUS AMPLIFICATION AND DETECTION OF SPECIFIC DNA SEQUENCES

Russell Higuchi*, Gavin Dollinger[1], P. Sean Walsh and Robert Griffith

Roche Molecular Systems, Inc., 1400 53rd St., Emeryville, CA 94608. [1]Chiron Corporation, 1400 53rd St., Emeryville, CA 94608. *Corresponding author.

We have enhanced the polymerase chain reaction (PCR) such that specific DNA sequences can be detected without opening the reaction tube. This enhancement requires the addition of ethidium bromide (EtBr) to a PCR. Since the fluorescence of EtBr increases in the presence of double-stranded (ds) DNA an increase in fluorescence in such a PCR indicates a positive amplification, which can be easily monitored externally. In fact, amplification can be continuously monitored in order to follow its progress. The ability to simultaneously amplify specific DNA sequences and detect the product of the amplification both simplifies and improves PCR and may facilitate its automation and more widespread use in the clinic or in other situations requiring high sample throughput.

Although the potential benefits of PCR[1] to clinical diagnostics are well known[2,3], it is still not widely used in this setting, even though it is four years since thermostable DNA polymerase[4] made PCR practical. Some of the reasons for its slow acceptance are high cost, lack of automation of pre- and post-PCR processing steps, and false positive results from carryover-contamination. The first two points are related in that labor is the largest contributor to cost at the present stage of PCR development. Most current assays require some form of "downstream" processing once thermocycling is done in order to determine whether the target DNA sequence was present and has amplified. These include DNA hybridization[5,6], gel electrophoresis with or without use of restriction digestion[7,8], HPLC[9], or capillary electrophoresis[10]. These methods are labor-intense, have low throughput, and are difficult to automate. The third point is also closely related to downstream processing. The handling of the PCR product in these downstream processes increases the chances that amplified DNA will spread through the typing lab, resulting in a risk of "carryover" false positives in subsequent testing[11].

These downstream processing steps would be eliminated if specific amplification and detection of amplified DNA took place simultaneously within an unopened reaction vessel. Assays in which such different processes take place without the need to separate reaction components have been termed "homogeneous". No truly homogeneous PCR assay has been demonstrated to date, although progress towards this end has been reported. Chehab, et al.[12], developed a PCR product detection scheme using fluorescent primers that resulted in a fluorescent PCR product. Allele-specific primers, each with different fluorescent tags, were used to indicate the genotype of the DNA. However, the unincorporated primers must still be removed in a downstream process in order to visualize the result. Recently, Holland, et al.[13], developed an assay in which the endogenous 5' exonuclease assay of *Taq* DNA polymerase was exploited to cleave a labeled oligonucleotide probe. The probe would only cleave if PCR amplification had produced its complementary sequence. In order to detect the cleavage products, however, a subsequent process is again needed.

We have developed a truly homogeneous assay for PCR and PCR product detection based upon the greatly increased fluorescence that ethidium bromide and other DNA binding dyes exhibit when they are bound to dsDNA[14–16]. As outlined in Figure 1, a prototypic PCR

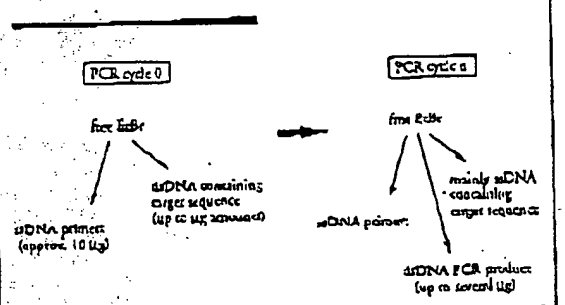


FIGURE 1 Principle of simultaneous amplification and detection of PCR product. The components of a PCR containing EtBr that are fluorescent are listed—EtBr itself, EtBr bound to either ssDNA or dsDNA. There is a large fluorescence enhancement when EtBr is bound to DNA and binding is greatly enhanced when DNA is double-stranded. After sufficient (n) cycles of PCR, the net increase in dsDNA results in additional EtBr binding, and a net increase in total fluorescence.

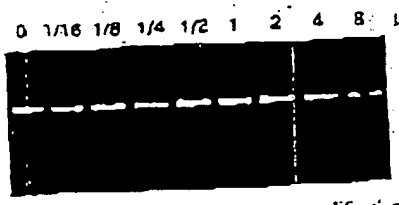


FIGURE 2 Gel electrophoresis of PCR amplification products of the human nuclear gene, HLA DQα, made in the presence of increasing amounts of EtBr (up to 8 µg/ml). The presence of EtBr has no obvious effect on the yield or specificity of amplification.

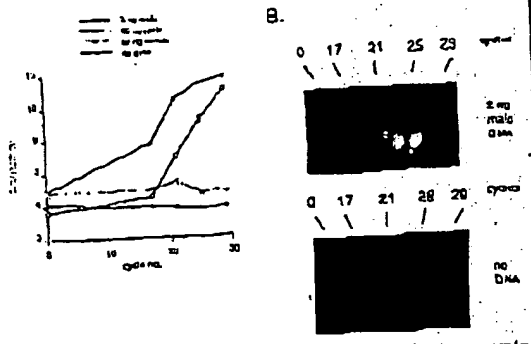


FIGURE 3 (A) Fluorescence measurements from PCRs that contain 0.5 µg/ml EtBr and that are specific for Y-chromosome repeat sequences. Five replicate PCRs were begun containing each of the DNAs specified. At each indicated cycle, one of the five replicate PCRs for each DNA was removed from thermocycling and its fluorescence measured. Units of fluorescence are arbitrary. (B) UV photography of PCR tubes (0.5 ml Eppendorf-style, polypropylene micro-centrifuge tubes) containing reactions, those starting from 2 ng male DNA and control reactions without any DNA, from (A).

begins with primers that are single-stranded DNA (ssDNA), dNTPs, and DNA polymerase. An amount of dsDNA containing the target sequence (target DNA) is also typically present. This amount can vary, depending on the application, from single-cell amounts of DNA[17] to micrograms per PCR[18]. If EtBr is present, the reagents that will fluoresce, in order of increasing fluorescence, are free EtBr itself, and EtBr bound to the single-stranded DNA primers and to the double-stranded target DNA (by its intercalation between the stacked bases of the DNA double-helix). After the first denaturation cycle, target DNA will be largely single-stranded. After a PCR is completed, the most significant change is the increase in the amount of dsDNA (the PCR product itself) of up to several micrograms. Formerly free EtBr is bound to the additional dsDNA, resulting in an increase in fluorescence. There is also some decrease in the amount of ssDNA primer, but because the binding of EtBr to ssDNA is much less than to dsDNA, the effect of this change on the total fluorescence of the sample is small. The fluorescence increase can be measured by directing excitation illumination through the walls of the amplification vessel before and after, or even continuously during, thermocycling.

RESULTS

PCR in the presence of EtBr. In order to assess the affect of EtBr in PCR, amplifications of the human HLA DQα gene[19] were performed with the dye present at concentrations from 0.06 to 8.0 µg/ml (a typical concentration of EtBr used in staining of nucleic acids following gel electrophoresis is 0.5 µg/ml). As shown in Figure 2, gel electrophoresis revealed little or no difference in the yield or quality of the amplification product whether EtBr was absent or present at any of these concentrations, indicating that EtBr does not inhibit PCR.

Detection of human Y-chromosome specific sequences. Sequence-specific, fluorescence enhancement of EtBr as a result of PCR was demonstrated in a series of amplifications containing 0.5 µg/ml EtBr and primers specific to repeat DNA sequences found on the human Y-chromosome[20]. These PCRs initially contained either 60 ng male, 60 ng female, 2 ng male human or no DNA. Five replicate PCRs were begun for each DNA. After 0, 17, 21, 24 and 29 cycles of thermocycling, a PCR for each DNA was removed from the thermocycler, and its fluorescence measured in a spectrofluorometer and plotted vs. amplification cycle number (Fig. 3A). The shape of this curve reflects the fact that by the time an increase in fluorescence can be detected, the increase in DNA is becoming linear and not exponential with cycle number. As shown, the fluorescence increased about three-fold over the background fluorescence for the PCRs containing human male DNA, but did not significantly increase for negative control PCRs, which contained either no DNA or human female DNA. The more male DNA present to begin with—60 ng versus 2 ng—the fewer cycles were needed to give a detectable increase in fluorescence. Gel electrophoresis on the products of these amplifications showed that DNA fragments of the expected size were made in the male DNA containing reactions and that little DNA synthesis took place in the control samples.

In addition, the increase in fluorescence was visualized by simply laying the completed, unopened PCRs on a UV transilluminator and photographing them through a red filter. This is shown in figure 3B for the reactions that began with 2 ng male DNA and those with no DNA.

Detection of specific alleles of the human β-globin gene. In order to demonstrate that this approach has adequate specificity to allow genetic screening, a detection of the sickle-cell anemia mutation was performed. Figure 4 shows the fluorescence from completed amplifications containing EtBr (0.5 µg/ml) as detected by photography of the reaction tubes on a UV transilluminator. These reactions were performed using primers specific for either the wild-type or sickle-cell mutation of the human β-globin gene[21]. The specificity for each allele is imparted by placing the sickle-mutation site at the terminal 3′ nucleotide of one primer. By using an appropriate primer annealing temperature, primer extension—and thus amplification—can take place only if the 3′ nucleotide of the primer is complementary to the β-globin allele present[21,22].

Each pair of amplifications shown in Figure 4 consists of a reaction with either the wild-type allele specific (left tube) or sickle-allele specific (right tube) primers. Three different DNAs were typed: DNA from a homozygous wild-type β-globin individual (AA); from a heterozygous sickle β-globin individual (AS); and from a homozygous sickle β-globin individual (SS). Each DNA (50 ng genomic DNA to start each PCR) was analyzed in triplicate (3 pairs

of reactions each). The DNA type was reflected in the relative fluorescence intensities in each pair of completed amplifications. There was a significant increase in fluorescence only where a β-globin allele DNA matched the primer set. When measured on a spectrofluorometer (data not shown), this fluorescence was about three times that present in a PCR where both β-globin alleles were mismatched to the primer set. Gel electrophoresis (not shown) established that this increase in fluorescence was due to the synthesis of nearly a microgram of a DNA fragment of the expected size for β-globin. There was little synthesis of dsDNA in reactions in which the allele-specific primer was mismatched to both alleles.

Continuous monitoring of a PCR. Using a fiber optic device, it is possible to direct excitation illumination from a spectrofluorometer to a PCR undergoing thermocycling and to return its fluorescence to the spectrofluorometer. The fluorescence readout of such an arrangement, directed at an EtBr-containing amplification of Y-chromosome specific sequences from 25 ng of human male DNA, is shown in Figure 5. The readout from a control PCR with no target DNA is also shown. Thirty cycles of PCR were monitored for each.

The fluorescence trace as a function of time clearly shows the effect of the thermocycling. Fluorescence intensity rises and falls inversely with temperature. The fluorescence intensity is minimum at the denaturation temperature (94°C) and maximum at the annealing/extension temperature (50°C). In the negative-control PCR, these fluorescence maxima and minima do not change significantly over the thirty thermocycles, indicating that there is little dsDNA synthesis without the appropriate target DNA, and there is little if any bleaching of EtBr during the continuous illumination of the sample.

In the PCR containing male DNA, the fluorescence maxima at the annealing/extension temperature begin to increase at about 4000 seconds of thermocycling, and continue to increase with time, indicating that dsDNA is being produced at a detectable level. Note that the fluorescence minima at the denaturation temperature do not significantly increase, presumably because at this temperature there is no dsDNA for EtBr to bind. Thus the course of the amplification is followed by tracking the fluorescence increase at the annealing temperature. Analysis of the products of these two amplifications by gel electrophoresis showed a DNA fragment of the expected size for the male DNA containing sample and no detectable DNA synthesis for the control sample.

DISCUSSION

Downstream processes such as hybridization to a sequence-specific probe can enhance the specificity of DNA detection by PCR. The elimination of these processes means that the specificity of this homogeneous assay depends solely on that of PCR. In the case of sickle-cell disease, we have shown that PCR alone has sufficient DNA sequence specificity to permit genetic screening. Using appropriate amplification conditions, there is little non-specific production of dsDNA in the absence of the appropriate target allele.

The specificity required to detect pathogens can be more or less than that required to do genetic screening, depending on the number of pathogens in the sample and the amount of other DNA that must be taken with the sample. A difficult target is HIV, which requires detection of a viral genome that can be at the level of a few copies per thousands of host cells[6]. Compared with genetic screening, which is performed on cells containing at least one copy of the target sequence, HIV detection requires both more specificity and the input of more total

FIGURE 4 UV photography of PCR tubes containing amplifications using EtBr that are specific to wild-type (A) or sickle (S) alleles of the human β-globin gene. The left of each pair of tubes contains allele-specific primers to the wild-type alleles, the right tube primers to the sickle allele. The photograph was taken after 30 cycles of PCR, and the input DNAs and the alleles they contain are indicated. Fifty ng of DNA was used to begin PCR. Typing was done in triplicate (3 pairs of PCRs) for each input DNA.

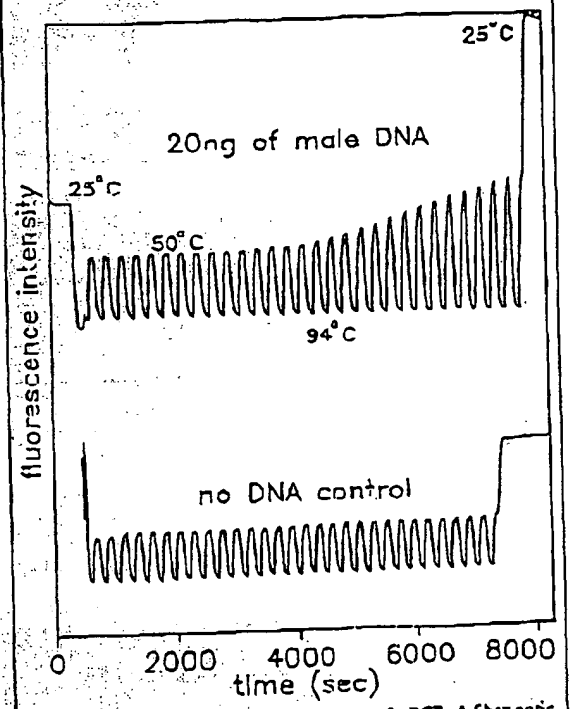


FIGURE 5 Continuous, real-time monitoring of a PCR. A fiber optic was used to carry excitation light to a PCR in progress and also emitted light back to a fluorometer (see Experimental Protocol). Amplification using human male-DNA specific primers in a PCR starting with 20 ng of human male DNA (top), or in a control PCR without DNA (bottom), were monitored. Thirty cycles of PCR were followed for each. The temperature cycled between 94°C (denaturation) and 50°C (annealing and extension). Note in the male DNA PCR, the cycle (time) dependent increase in fluorescence at the annealing/extension temperature.

DNA—up to microgram amounts—in order to have sufficient numbers of target sequences. This large amount of starting DNA in an amplification significantly increases the background fluorescence over which any additional fluorescence produced by PCR must be detected. An additional complication that occurs with targets in low copy-number is the formation of the "primer-dimer" artifact. This is the result of the extension of one primer using the other primer as a template. Although this occurs infrequently, once it occurs the extension product is a substrate for PCR amplification, and can compete with true PCR targets if those targets are rare. The primer-dimer product is of course dsDNA and thus is a potential source of false signal in this homogeneous assay.

To increase PCR specificity and reduce the effect of primer-dimer amplification, we are investigating a number of approaches, including the use of nested-primer amplifications that take place in a single tube[3], and the "hot-start", in which nonspecific amplification is reduced by raising the temperature of the reaction before DNA synthesis begins[26]. Preliminary results using these approaches suggest that primer-dimer is effectively reduced and it is possible to detect the increase in EtBr fluorescence in a PCR instigated by a single HIV genome in a background of 10^5 cells. With larger numbers of cells, the background fluorescence contributed by genomic DNA becomes problematic. To reduce this background, it may be possible to use sequence-specific DNA-binding dyes that can be made to preferentially bind PCR product over genomic DNA by incorporating the dye-binding DNA sequence into the PCR product through a 5' "add-on" to the oligonucleotide primer[24].

We have shown that the detection of fluorescence generated by an EtBr-containing PCR is straightforward, both once PCR is completed and continuously during thermocycling. The ease with which automation of specific DNA detection can be accomplished is the most promising aspect of this assay. The fluorescence analysis of completed PCRs is already possible with existing instrumentation in 96-well format[25]. In this format, the fluorescence in each PCR can be quantitated before, after, and even at selected points during thermocycling by moving the rack of PCRs to a 96-microwell plate fluorescence reader[26].

The instrumentation necessary to continuously monitor multiple PCRs simultaneously is also simple in principle. A direct extension of the apparatus used here is to have multiple fiberoptics transmit the excitation light and fluorescent emissions to and from multiple PCRs. The ability to monitor multiple PCRs continuously may allow quantitation of target DNA copy number. Figure 3 shows that the larger the amount of starting target DNA, the sooner during PCR a fluorescence increase is detected. Preliminary experiments (Higuchi and Dollinger, manuscript in preparation) with continuous monitoring have shown a sensitivity to two-fold differences in initial target DNA concentration.

Conversely, if the number of target molecules is known—as it can be in genetic screening—continuous monitoring may provide a means of detecting false positive and false negative results. With a known number of target molecules, a true positive would exhibit detectable fluorescence by a predictable number of cycles of PCR. Increases in fluorescence detected before or after that cycle would indicate potential artifacts. False negative results due to, for example, inhibition of DNA polymerase, may be detected by including within each PCR an inefficiently amplifying marker. This marker results in a fluorescence increase only after a large number of cycles—many more than are necessary to detect a true positive. If a sample fails to have a fluorescence increase after this many cycles, inhibition may be suspected. Since, in this assay, conclusions are drawn based on the presence or absence of fluorescence signal alone, such controls may be important. In any event, before any test based on this principle is ready for the clinic, an assessment of its false positive/false negative rates will need to be obtained using a large number of known samples.

In summary, the inclusion in PCR of dyes whose fluorescence is enhanced upon binding dsDNA makes it possible to detect specific DNA amplification from outside the PCR tube. In the future, instruments based upon this principle may facilitate the more widespread use of PCR in applications that demand the high throughput of samples.

EXPERIMENTAL PROTOCOL

Human HLA-DQα gene amplifications containing EtBr. PCRs were set up in 100 µl volumes containing 10 mM Tris-HCl, pH 8.3; 50 mM KCl; 4 mM $MgCl_2$; 2.5 units of *Taq* DNA polymerase (Perkin-Elmer Cetus, Norwalk, CT); 20 pmole each of human HLA-DQα gene specific oligonucleotide primers GH26 and GH27[15] and approximately 10^7 copies of DQα PCR product diluted from a previous reaction. Ethidium bromide (EtBr; Sigma) was used at the concentrations indicated in Figure 2. Thermocycling proceeded for 20 cycles in a model 480 thermocycler (Perkin-Elmer Cetus, Norwalk, CT) using a "step-cycle" program of 94°C for 1 min. denaturation and 60°C for 90 sec. annealing and 72°C for 30 sec. extension.

Y-chromosome specific PCR. PCRs (100 µl total reaction volume) containing 0.5 µg/ml EtBr were prepared as described for HLA-DQα, except with different primers and target DNAs. These PCRs contained 15 pmole each male DNA-specific primers Y1.1 and Y1.2[28], and either 50 ng male, 50 ng female, 2 ng male, or no human DNA. Thermocycling was 94°C for 1 min. and 60°C for 1 min using a "step-cycle" program. The number of cycles for a sample were as indicated in Figure 3. Fluorescence measurement is described below.

Allele-specific, human β-globin gene PCR. Amplifications of 100 µl volume using 0.5 µg/ml of EtBr were prepared as described for HLA-DQα above except with different primers and target DNAs. These PCRs contained either primer pair HGP2/HB14A (wild-type globin specific primers) or HGP2/HB14S (sickle-globin specific primers) at 10 pmole each primer per PCR. These primers were developed by Wu et al.[21]. Three different target DNAs were used in separate amplifications—50 ng each of human DNA that was homozygous for the sickle trait (SS), DNA that was heterozygous for the sickle trait (AS), or DNA that was homozygous for the wild globin (AA). Thermocycling was for 30 cycles at 94°C for 1 min. and 55°C for 1 min. using a "step-cycle" program. An annealing temperature of 55°C had been shown by Wu et al.[21] to provide allele-specific amplification. Completed PCRs were photographed through a red filter (Wratten 88A) after placing the reaction tubes atop a model TM-36 transilluminator (UV-products San Gabriel, CA).

Fluorescence measurement. Fluorescence measurements were made on PCRs containing EtBr in a Fluorolog-2 fluorometer (SPEX, Edison, NJ). Excitation was at the 500 nm band with about 9 nm bandwidth with a GG 435 nm cut-off filter (Melles Griot, Inc., Irvine, CA) to exclude second-order light. Emitted light was detected at 570 nm with a bandwidth of about 7 nm. An OG 530 nm cut-off filter was used to remove the excitation light.

Continuous fluorescence monitoring of PCR. Continuous monitoring of a PCR in progress was accomplished using the spectrofluorometer and settings described above as well as a fiberoptic accessory (SPEX cat. no. 1950) to both send excitation light to, and receive emitted light from, a PCR placed in a well of a model 480 thermocycler (Perkin-Elmer Cetus). The probe end of the fiberoptic cable was attached with "5 minute-epoxy" to the open top of a PCR tube (a 0.5 ml polypropylene centrifuge tube with its cap removed) effectively sealing it. The exposed top of the PCR tube and the end of the fiberoptic cable were shielded from room light and the room lights were kept dimmed during each run. The monitored PCR was an amplification of Y-chromosome-specific repeat sequences as described above, except using an annealing/extension temperature of 50°C. The reaction was covered with mineral oil (2 drops) to prevent evaporation. Thermocycling and fluorescence measurement were started simultaneously. A time-base scan with a 10 second integration

was used and the emission signal was ratioed to the excitation signal to control for changes in light-source intensity. Data were collected using the dm3000f, version 2.5 (SPEX) data system.

Acknowledgments

We thank Bob Jones for help with the spectrofluorometric measurements and Heatherbell Fong for editing this manuscript.

References

1. Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. 1986. Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. CSHSQB 51:263–273.
2. White, T. J., Arnheim, N. and Erlich, H. A. 1989. The polymerase chain reaction. Trends Genet. 5:185–189.
3. Erlich, H. A., Gelfand, D. and Sninsky, J. J. 1991. Recent advances in the polymerase chain reaction. Science 252:1643–1651.
4. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491.
5. Saiki, R. K., Walsh, P. S., Levenson, C. H. and Erlich, H. A. 1989. Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. Proc. Natl. Acad. Sci. USA 86:6230–6234.
6. Kwok, S. Y., Mack, D. H., Mullis, K. B., Poiesz, B. J., Ehrlich, G. D., Blair, D. and Friedman-Kien, A. E. 1987. Identification of human immunodeficiency virus sequences by using in vitro enzymatic amplification and oligomer cleavage detection. J. Virol. 61:1690–1694.
7. Chehab, F. F., Doherty, M., Cai, S. P., Kan, Y. W., Cooper, S. and Rubin, E. M. 1987. Detection of sickle cell anemia and thalassemias. Nature 329:293–294.
8. Horn, G. T., Richards, B. and Klinger, K. W. 1989. Amplification of a highly polymorphic VNTR segment by the polymerase chain reaction. Nuc. Acids Res. 16:2140.
9. Katz, E. D. and Dong, M. W. 1990. Rapid analysis and purification of polymerase chain reaction products by high-performance liquid chromatography. Biotechniques 8:546–555.
10. Heiger, D. N., Cohen, A. S. and Karger, B. L. 1990. Separation of DNA restriction fragments by high performance capillary electrophoresis with low and zero crosslinked polyacrylamide using continuous and pulsed electric fields. J. Chromatogr. 516:33–48.
11. Kwok, S. Y. and Higuchi, R. G. 1989. Avoiding false positives with PCR. Nature 339:237–238.
12. Chehab, F. F. and Kan, Y. W. 1989. Detection of specific DNA sequences by fluorescence amplification: a color complementation assay. Proc. Natl. Acad. Sci. USA 86:9178–9182.
13. Holland, P. M., Abramson, R. D., Watson, R. and Gelfand, D. H. 1991. Detection of specific polymerase chain reaction product by utilizing the 5′ to 3′ exonuclease activity of *Thermus aquaticus* DNA polymerase. Proc. Natl. Acad. Sci. USA 88:7276–7280.
14. Markovits, J., Roques, B. P. and Le Pecq, J. B. 1979. Ethidium dimer: a new reagent for the fluorimetric determination of nucleic acids. Anal. Biochem. 94:259–264.
15. Kapuscinski, J. and Szer, W. 1979. Interactions of 4′,6-diamidine-2-phenylindole with synthetic polynucleotides. Nuc. Acids Res. 6:3519–3534.
16. Searle, M. S. and Embrey, K. J. 1990. Sequence-specific interaction of Hoechst 33258 with the minor groove of an adenine-tract DNA duplex studied in solution by ^1H NMR spectroscopy. Nuc. Acids Res. 18:3753–3762.
17. Li, H. H., Gyllensten, U. B., Cui, X. F., Saiki, R. K., Erlich, H. A. and Arnheim, N. 1988. Amplification and analysis of DNA sequences in single human sperm and diploid cells. Nature 335:414–417.
18. Abbott, M. A., Poiesz, B. J., Byrne, B. C., Kwok, S. Y., Sninsky, J. J. and Erlich, H. A. 1988. Enzymatic gene amplification: qualitative and quantitative methods for detecting proviral DNA amplified in vitro. J. Infect. Dis. 158:1158.
19. Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B. and Erlich, H. A. 1986. Analysis of enzymatically amplified β-globin and HLA-DQα DNA with allele-specific oligonucleotide probes. Nature 324:163–166.
20. Kogan, S. C., Doherty, M. and Gitschier, J. 1987. An improved method for prenatal diagnosis of genetic diseases by analysis of amplified DNA sequences. N. Engl. J. Med. 317:985–990.
21. Wu, D. Y., Ugozzoli, L., Pal, B. K. and Wallace, R. B. 1989. Allele-specific enzymatic amplification of β-globin genomic DNA for diagnosis of sickle cell anemia. Proc. Natl. Acad. Sci. USA 86:2757–2760.
22. Kwok, S., Kellogg, D. E., McKinney, N., Spasic, D., Goda, L., Levenson, C. and Sninsky, J. J. 1990. Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies. Nuc. Acids Res. 18:999–1005.
23. Chou, Q., Russell, M., Birch, D., Raymond, J. and Bloch, W. 1992. Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. Submitted.
24. Higuchi, R. 1989. Using PCR to engineer DNA, p. 61–70. *In:* PCR Technology. H. A. Erlich (Ed.). Stockton Press, New York, N.Y.
25. Haff, L., Atwood, J. G., DiCesare, J., Katz, E., Picozza, E., Williams, J. F. and Woudenberg, T. 1991. A high-performance system for automation of the polymerase chain reaction. Biotechniques 10:102–103, 106–112.
26. Tumosa, N. and Kabat, L. 1988. Fluorescent EIA screening of monoclonal antibodies to cell surface antigens. J. Immun. Meth. 116:59–65.

HellerEhrman
ATTORNEYS

Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 14717–14722, December 1998
Cell Biology, Medical Sciences

WISP genes are members of the connective tissue growth factor family that are up-regulated in Wnt-1-transformed cells and aberrantly expressed in human colon tumors

DIANE PENNICA*†, TODD A. SWANSON*, JAMES W. WELSH*, MARGARET A. ROY‡, DAVID A. LAWRENCE*, JAMES LEE‡, JENNIFER BRUSH‡, LISA A. TANEYHILL§, BETHANNE DEUEL‡, MICHAEL LEW¶, COLIN WATANABE‖, ROBERT L. COHEN*, MONA P. MELHEM**, GENE G. FINLEY**, PHIL QUIRKE††, AUDREY D. GODDARD‡, KENNETH J. HILLAN¶, AUSTIN L. GURNEY‡, DAVID BOTSTEIN‡‡, AND ARNOLD J. LEVINE§

Departments of *Molecular Oncology, ‡Molecular Biology, Scientific Computing, and **Pathology, Genentech Inc., 1 DNA Way, South San Francisco, CA 94080; **University of Pittsburgh School of Medicine, Veterans Administration Medical Center, Pittsburgh, PA 15240; ††University of Leeds, Leeds, LS2 9JT United Kingdom; ‡‡Department of Genetics, Stanford University, Palo Alto, CA 94305; and §Department of Molecular Biology, Princeton University, Princeton, NJ 08544

Contributed by David Botstein and Arnold J. Levine, October 21, 1998

ABSTRACT Wnt family members are critical to many developmental processes, and components of the Wnt signaling pathway have been linked to tumorigenesis in familial and sporadic colon carcinomas. Here we report the identification of two genes, *WISP-1* and *WISP-2*, that are up-regulated in the mouse mammary epithelial cell line CS7MG transformed by Wnt-1, but not by Wnt-4. Together with a third related gene, *WISP-3*, these proteins define a subfamily of the connective tissue growth factor family. Two distinct systems demonstrated *WISP* induction to be associated with the expression of Wnt-1. These included (*i*) CS7MG cells infected with a Wnt-1 retroviral vector or expressing Wnt-1 under the control of a tetracycline repressible promoter, and (*ii*) Wnt-1 transgenic mice. The *WISP-1* gene was localized to human chromosome 8q24.1–8q24.3. *WISP-1* genomic DNA was amplified in colon cancer cell lines and in human colon tumors and its RNA overexpressed (2- to >30-fold) in 84% of the tumors examined compared with patient-matched normal mucosa. *WISP-3* mapped to chromosome 6q22–6q23 and also was overexpressed (4- to >40-fold) in 63% of the colon tumors analyzed. In contrast, *WISP-2* mapped to human chromosome 20q12–20q13 and its DNA was amplified, but RNA expression was reduced (2- to >30-fold) in 79% of the tumors. These results suggest that the *WISP* genes may be downstream of Wnt-1 signaling and that aberrant levels of *WISP* expression in colon cancer may play a role in colon tumorigenesis.

Wnt-1 is a member of an expanding family of cysteine-rich, glycosylated signaling proteins that mediate diverse developmental processes such as the control of cell proliferation, adhesion, cell polarity, and the establishment of cell fates (1, 2). Wnt-1 originally was identified as an oncogene activated by the insertion of mouse mammary tumor virus in virus-induced mammary adenocarcinomas (3, 4). Although Wnt-1 is not expressed in the normal mammary gland, expression of Wnt-1 in transgenic mice causes mammary tumors (5).

In mammalian cells, Wnt family members initiate signaling by binding to the seven-transmembrane spanning Frizzled receptors and recruiting the cytoplasmic protein Dishevelled (Dsh) to the cell membrane (1, 2, 6). Dsh then inhibits the kinase activity of the normally constitutively active glycogen synthase kinase-3β (GSK-3β) resulting in an increase in β-catenin levels. Stabilized β-catenin interacts with the transcription factor TCF/Lef1, forming a complex that appears in the nucleus and binds TCF/Lef1 target DNA elements to activate transcription (7, 8). Other experiments suggest that the adenomatous polyposis coli (APC) tumor suppressor gene also plays an important role in Wnt signaling by regulating β-catenin levels (9). APC is phosphorylated by GSK-3β, binds to β-catenin, and facilitates its degradation. Mutations in either APC or β-catenin have been associated with colon carcinomas and melanomas, suggesting these mutations contribute to the development of these types of cancer, implicating the Wnt pathway in tumorigenesis (1).

Although much has been learned about the Wnt signaling pathway over the past several years, only a few of the transcriptionally activated downstream components activated by Wnt have been characterized. Those that have been described cannot account for all of the diverse functions attributed to Wnt signaling. Among the candidate Wnt target genes are those encoding the nodal-related 3 gene, *Xnr3*, a member of the transforming growth factor (TGF)-β superfamily, and the homeobox genes, *engrailed*, *goosecoid*, *twin* (*Xtwn*), and *siamois* (2). A recent report also identifies *c-myc* as a target gene of the Wnt signaling pathway (10).

To identify additional downstream genes in the Wnt signaling pathway that are relevant to the transformed cell phenotype, we used a PCR-based cDNA subtraction strategy, suppression subtractive hybridization (SSH) (11), using RNA isolated from C57MG mouse mammary epithelial cells and C57MG cells stably transformed by a Wnt-1 retrovirus. Overexpression of Wnt-1 in this cell line is sufficient to induce a partially transformed phenotype, characterized by elongated and refractile cells that lose contact inhibition and form a multilayered array (12, 13). We reasoned that genes differentially expressed between these two cell lines might contribute to the transformed phenotype.

In this paper, we describe the cloning and characterization of two genes up-regulated in Wnt-1 transformed cells, *WISP-1* and *WISP-2*, and a third related gene, *WISP-3*. The *WISP* genes are members of the CCN family of growth factors, which includes connective tissue growth factor (CTGF), Cyr61, and *nov*, a family not previously linked to Wnt signaling.

MATERIALS AND METHODS

SSH. SSH was performed by using the PCR-Select cDNA Subtraction Kit (CLONTECH). Tester double-stranded

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: TGF, transforming growth factor; CTGF, connective tissue growth factor; SSH, suppression subtractive hybridization; VWC, von Willebrand factor type C module.

Data deposition: The sequences reported in this paper have been deposited in the Genbank database (accession nos. AF100777, AF100778, AF100779, AF100780, and AF100781).

†To whom reprint requests should be addressed. e-mail: diane@genec.com.

cDNA was synthesized from 2 μg of poly(A)$^+$ RNA isolated from the CS7MG/Wnt-1 cell line and driver cDNA from 2 μg of poly(A)$^+$ RNA from the parent CS7MG cells. The subtracted cDNA library was subcloned into a pGEM-T vector for further analysis.

cDNA Library Screening. Clones encoding full-length mouse *WISP-1* were isolated by screening a λgt10 mouse embryo cDNA library (CLONTECH) with a 711-bp probe from the original partial clone S68 sequence corresponding to amino acids 128–169. Clones encoding full-length human *WISP-1* were isolated by screening λgt10 lung and fetal kidney cDNA libraries with the same probe at low stringency. Clones encoding full-length mouse and human *WISP-2* were isolated by screening a CS7MG/Wnt-1 or human fetal lung cDNA library with a probe corresponding to nucleotides 1463–1512. Full-length cDNAs encoding *WISP-3* were cloned from human bone marrow and fetal kidney libraries.

Expression of Human *WISP* RNA. PCR amplification of first-strand cDNA was performed with human Multiple Tissue cDNA panels (CLONTECH) and 300 μM of each dNTP at 94°C for 1 sec, 62°C for 30 sec, 72°C for 1 min, for 22–32 cycles. *WISP* and glyceraldehyde-3-phosphate dehydrogenase primer sequences are available on request.

***In Situ* Hybridization.** ^{35}P-labeled sense and antisense riboprobes were transcribed from an 897-bp PCR product corresponding to nucleotides 601–1440 of mouse *WISP-1* or a 294-bp PCR product corresponding to nucleotides 62–375 of mouse *WISP-2*. All tissues were processed as described (40).

Radiation Hybrid Mapping. Genomic DNA from each hybrid in the Stanford G3 and Genebridge4 Radiation Hybrid Panels (Research Genetics, Huntsville, AL) and human and hamster control DNAs were PCR-amplified, and the results were submitted to the Stanford or Massachusetts Institute of Technology web servers.

Cell Lines, Tumors, and Mucosa Specimens. Tissue specimens were obtained from the Department of Pathology (University of Pittsburgh) for patients undergoing colon resection and from the University of Leeds, United Kingdom. Genomic DNA was isolated (Qiagen) from the pooled blood of 10 normal human donors, surgical specimens, and the following ATCC human cell lines: SW480, COLO 320DM, HT-29, WiDr, and SW403 (colon adenocarcinomas), SW620 (lymph node metastasis, colon adenocarcinoma), HCT 116 (colon carcinoma), SK-CO-1 (colon adenocarcinoma, ascites), and HM7 (a variant of ATCC colon adenocarcinoma cell line LS 174T). DNA concentration was determined by using Hoechst dye 33258 intercalation fluorimetry. Total RNA was prepared by homogenization in 7 M GuSCN followed by centrifugation over CsCl cushions or prepared by using RNAzol.

Gene Amplification and RNA Expression Analysis. Relative gene amplification and RNA expression of *WISPs* and *c-myc* in the cell lines, colorectal tumors, and normal mucosa were determined by quantitative PCR. Gene-specific primers and fluorogenic probes (sequences available on request) were designed and used to amplify and quantitate the genes. The relative gene copy number was derived by using the formula $2^{(\Delta Ct)}$ where ΔCt represents the difference in amplification cycles required to detect the *WISP* genes in peripheral blood lymphocyte DNA compared with colon tumor DNA or colon tumor RNA compared with normal mucosal RNA. The δ-method was used for calculation of the SE of the gene copy number or RNA expression level. The *WISP*-specific signal was normalized to that of the glyceraldehyde-3-phosphate dehydrogenase housekeeping gene. All TaqMan assay reagents were obtained from Perkin–Elmer Applied Biosystems.

RESULTS

Isolation of *WISP-1* and *WISP-2* by SSH. To identify Wnt-1-inducible genes, we used the technique of SSH using the mouse mammary epithelial cell line CS7MG and CS7MG cells that stably express Wnt-1 (11). Candidate differentially expressed cDNAs (1,384 total) were sequenced. Thirty-nine percent of the sequences matched known genes or homologues, 32% matched expressed sequence tags, and 29% had no match. To confirm that the transcript was differentially expressed, semiquantitative reverse transcription–PCR and Northern analysis were performed by using mRNA from the CS7MG and CS7MG/Wnt-1 cells.

Two of the cDNAs, *WISP-1* and *WISP-2*, were differentially expressed, being induced in the CS7MG/Wnt-1 cell line, but not in the parent CS7MG cells or CS7MG cells overexpressing Wnt-4 (Fig. 1 *A* and *B*). Wnt-4, unlike Wnt-1, does not induce the morphological transformation of CS7MG cells and has no effect on β-catenin levels (13, 14). Expression of *WISP-1* was up-regulated approximately 3-fold in the CS7MG/Wnt-1 cell line and *WISP-2* by approximately 5-fold by both Northern analysis and reverse transcription–PCR.

An independent, but similar, system was used to examine *WISP* expression after Wnt-1 induction. CS7MG cells expressing the *Wnt-1* gene under the control of a tetracycline-repressible promoter produce low amounts of Wnt-1 in the repressed state but show a strong induction of *Wnt-1* mRNA and protein within 24 hr after tetracycline removal (8). The levels of Wnt-1 and *WISP* RNA isolated from these cells at various times after tetracycline removal were assessed by quantitative PCR. Strong induction of Wnt-1 mRNA was seen as early as 10 hr after tetracycline removal. Induction of *WISP* mRNA (2- to 6-fold) was seen at 48 and 72 hr (data not shown). These data support our previous observations that show that *WISP* induction is correlated with Wnt-1 expression. Because the induction is slow, occurring after approximately 48 hr, the induction of *WISPs* may be an indirect response to Wnt-1 signaling.

cDNA clones of human *WISP-1* were isolated and the sequence compared with mouse *WISP-1*. The cDNA sequences of mouse and human *WISP-1* were 1,766 and 2,830 bp in length, respectively, and encode proteins of 367 aa, with predicted relative molecular masses of ≈40,000 (M_r 40 K). Both have hydrophobic N-terminal signal sequences, 38 conserved cysteine residues, and four potential N-linked glycosylation sites and are 84% identical (Fig. 2*A*).

Full-length cDNA clones of mouse and human *WISP-2* were 1,734 and 1,293 bp in length, respectively, and encode proteins of 251 and 250 aa, respectively, with predicted relative molecular masses of ≈27,000 (M_r 27 K) (Fig. 2*B*). Mouse and human *WISP-2* are 73% identical. Human *WISP-2* has no potential N-linked glycosylation sites, and mouse *WISP-2* has one at

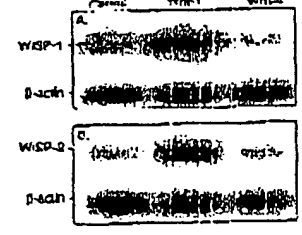


FIG. 1. *WISP-1* and *WISP-2* are induced by Wnt-1, but not Wnt-4, expression in CS7MG cells. Northern analysis of *WISP-1* (*A*) and *WISP-2* (*B*) expression in CS7MG, CS7MG/Wnt-1, and CS7MG/Wnt-4 cells. Poly(A)$^+$ RNA (2 μg) was subjected to Northern blot analysis and hybridized with a 70-bp mouse *WISP-1*-specific probe (amino acids 278–300) or a 190-bp *WISP-2*-specific probe (nucleotides 1438–1627) in the 3′ untranslated region. Blots were rehybridized with human β-actin probe.

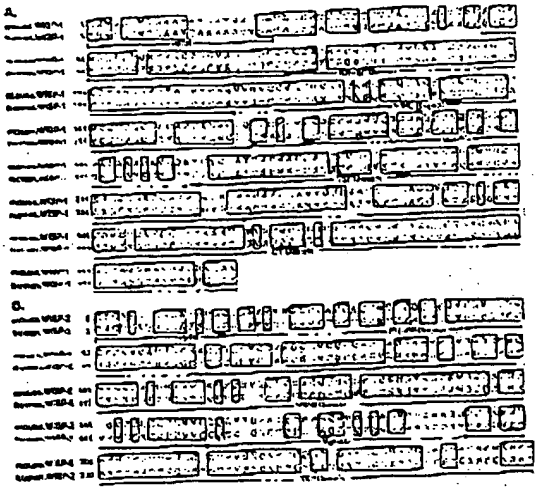


FIG. 2. Encoded amino acid sequence alignment of mouse and human *WISP-1* (*A*) and mouse and human *WISP-2* (*B*). The potential signal sequence, insulin-like growth factor-binding protein (IGF-BP), VWC, thrombospondin (TSP), and C-terminal (CT) domains are underlined.

position 197. *WISP-2* has 28 cysteine residues that are conserved among the 38 cysteines found in *WISP-1*.

Identification of *WISP-3*. To search for related proteins, we screened expressed sequence tag (EST) databases with the WISP-1 protein sequence and identified several ESTs as potentially related sequences. We identified a homologous protein that we have called WISP-3. A full-length human *WISP-3* cDNA of 1,371 bp was isolated corresponding to those ESTs that encode a 354-aa protein with a predicted molecular mass of 39,293. WISP-3 has two potential N-linked glycosylation sites and 36 cysteine residues. An alignment of the three human WISP proteins shows that WISP-1 and WISP-3 are the most similar (42% identity), whereas WISP-2 has 37% identity with WISP-1 and 32% identity with WISP-3 (Fig. 3*A*).

***WISPs* Are Homologous to the CTGF Family of Proteins.** Human *WISP-1*, *WISP-2*, and *WISP-3* are novel sequences; however, mouse *WISP-1* is the same as the recently identified *Elm1* gene. *Elm1* is expressed in low, but not high, metastatic mouse melanoma cells, and suppresses the *in vivo* growth and metastatic potential of K-1735 mouse melanoma cells (15). Human and mouse *WISP-2* are homologous to the recently described rat gene, *rCop-1* (16). Significant homology (36–44%) was seen to the CCN family of growth factors. This family includes three members, CTGF, Cyr61, and the protooncogene *nov*. CTGF is a chemotactic and mitogenic factor for fibroblasts that is implicated in wound healing and fibrotic disorders and is induced by TGF-*β* (17). Cyr61 is an extracellular matrix signaling molecule that promotes cell adhesion, proliferation, migration, angiogenesis, and tumor growth (18, 19). *nov* (nephroblastoma overexpressed) is an immediate early gene associated with quiescence and found altered in Wilms tumors (20). The proteins of the CCN family share functional, but not sequence, similarity to Wnt-1. All are secreted, cysteine-rich heparin binding glycoproteins that associate with the cell surface and extracellular matrix.

WISP proteins exhibit the modular architecture of the CCN family, characterized by four conserved cysteine-rich domains (Fig. 3*B*) (21). The N-terminal domain, which includes the first 12 cysteine residues, contains a consensus sequence (GCGC-CXXC) conserved in most insulin-like growth factor (IGF)-binding proteins (BP). This sequence is conserved in WISP-2 and WISP-3, whereas WISP-1 has a glutamine in the third position instead of a glycine. CTGF recently has been shown to specifically bind IGF (22) and a truncated *nov* protein lacking the IGF-BP domain is oncogenic (23). The von Willebrand factor type C module (VWC), also found in certain collagens and mucins, covers the next 10 cysteine residues, and is thought to participate in protein complex formation and oligomerization (24). The VWC domain of WISP-3 differs from all CCN family members described previously, in that it contains only six of the 10 cysteine residues (Fig. 3 *A* and *B*). A short variable region follows the VWC domain. The third module, the thrombospondin (TSP) domain is involved in binding to sulfated glycoconjugates and contains six cysteine residues and a conserved W5xCSxxCG motif first identified in thrombospondin (25). The C-terminal (CT) module containing the remaining 10 cysteines is thought to be involved in dimerization and receptor binding (26). The CT domain is present in all CCN family members described to date but is absent in WISP-2 (Fig. 3 *A* and *B*). The existence of a putative signal sequence and the absence of a transmembrane domain suggest that WISPs are secreted proteins, an observation supported by an analysis of their expression and secretion from mammalian cell and baculovirus cultures (data not shown).

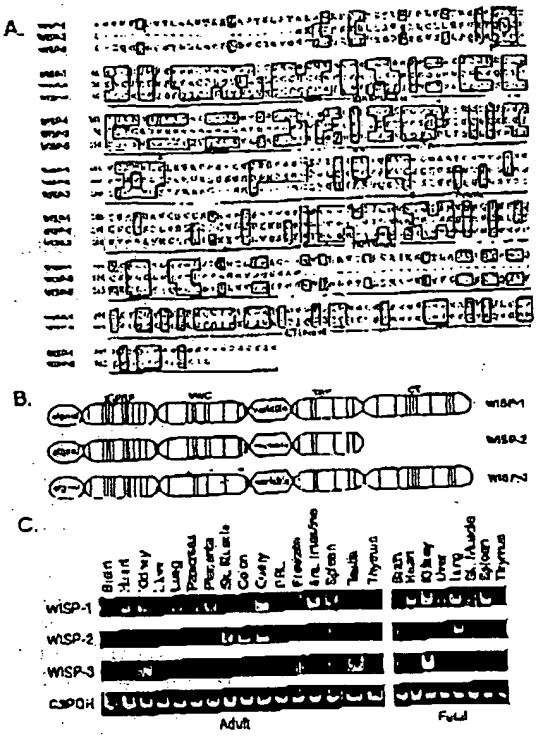


FIG. 3. (*A*) Encoded amino acid sequence alignment of human WISPs. The cysteine residues of WISP-1 and WISP-2 that are not present in WISP-3 are indicated with a dot. (*B*) Schematic representation of the WISP proteins showing the domain structure and cysteine residues (vertical lines). The four cysteine residues in the VWC domain that are absent in WISP-3 are indicated with a dot. (*C*) Expression of *WISP* mRNA in human tissues. PCR was performed on human multiple-tissue cDNA panels (CLONTECH) from the indicated adult and fetal tissues.

Expression of *WISP* mRNA in Human Tissues. Tissue-specific expression of human *WISPs* was characterized by PCR

analysis on adult and fetal multiple tissue cDNA panels. *WISP-1* expression was seen in the adult heart, kidney, lung, pancreas, placenta, ovary, small intestine, and spleen (Fig. 3*C*). Little or no expression was detected in the brain, liver, skeletal muscle, colon, peripheral blood leukocytes, prostate, testis, or thymus. *WISP-2* had a more restricted tissue expression and was detected in adult skeletal muscle, colon, ovary, and fetal lung. Predominant expression of *WISP-3* was seen in adult kidney and testis and fetal kidney. Lower levels of *WISP-3* expression were detected in placenta, ovary, prostate, and small intestine.

***In Situ* Localization of *WISP-1* and *WISP-2*.** Expression of *WISP-1* and *WISP-2* was assessed by *in situ* hybridization in mammary tumors from Wnt-1 transgenic mice. Strong expression of *WISP-1* was observed in stromal fibroblasts lying within the fibrovascular tumor stroma (Fig. 4 *A–D*). However, low-level *WISP-1* expression also was observed focally within tumor cells (data not shown). No expression was observed in normal breast. Like *WISP-1*, *WISP-2* expression also was seen in the tumor stroma in breast tumors from Wnt-1 transgenic animals (Fig. 4 *E–H*). However, *WISP-2* expression in the stroma was in spindle-shaped cells adjacent to capillary vessels, whereas

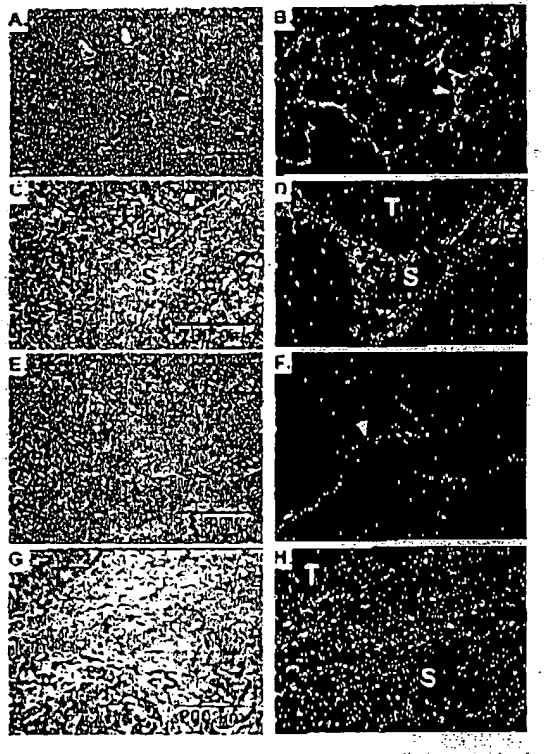


FIG. 4. (*A*, *C*, *E*, and *G*) Representative hematoxylin/eosin-stained images from breast tumors in Wnt-1 transgenic mice. The corresponding dark-field images showing *WISP-1* expression are shown in *B* and *D*. The tumor is a moderately well-differentiated adenocarcinoma showing evidence of adenoid cystic change. At low power (*A* and *B*), expression of *WISP-1* is seen in the delicate branching fibrovascular tumor stroma (arrowhead). At higher magnification, expression is seen in the stromal(s) fibroblasts (*C* and *D*), and tumor cells are negative. Focal expression of *WISP-1*, however, was observed in tumor cells in some areas. Images of *WISP-2* expression are shown in *E–H*. At low power (*E* and *F*), expression of WISP-2 is seen in cells lying within the fibrovascular tumor stroma. At higher magnification, these cells appeared to be adjacent to capillary vessels whereas tumor cells are negative (*G* and *H*).

the predominant cell type expressing *WISP-1* was the stromal fibroblasts.

Chromosome Localization of the *WISP* Genes. The chromosomal location of the human *WISP* genes was determined by radiation hybrid mapping panels. *WISP-1* is approximately 3.48 cR from the meiotic marker AFM259xc5 [logarithm of odds (lod) score 16.31] on chromosome 8q24.1 to 8q24.3. In the same region as the human locus of the *nov*H family member (27) and roughly 4 Mbs distal to *c-myc* (28). Preliminary fine mapping indicates that *WISP-1* is located near D8S1712 STS. *WISP-2* is linked to the marker SHGC-33922 (lod = 1,000) on chromosome 20q12–20q13.1. Human *WISP-3* mapped to chromosome 6q22–6q23 and is linked to the marker AFM211xe5 (lod = 1,000). *WISP-3* is approximately 13 Mbs proximal to CTGF and 23 Mbs proximal to the human cellular oncogene *MYB* (27, 29).

Amplification and Aberrant Expression of *WISPs* in Human Colon Tumors. Amplification of protooncogenes is seen in many human tumors and has etiological and prognostic significance. For example, in a variety of tumor types, *c-myc* amplification has been associated with malignant progression and poor prognosis (30). Because *WISP-1* resides in the same general chromosomal location (8q24) as *c-myc*, we asked whether it was a target of gene amplification, and, if so, whether this amplification was independent of the *c-myc* locus. Genomic DNA from human colon cancer cell lines was assessed by quantitative PCR and Southern blot analysis. (Fig. 5 *A* and *B*). Both methods detected similar degrees of *WISP-1* amplification. Most cell lines showed significant (2- to 4-fold) amplification, with the HT-29 and WiDr cell lines demonstrating an 8-fold increase. Significantly, the pattern of amplification observed did not correlate with that observed for *c-myc*, indicating that the *c-myc* gene is not part of the amplicon that involves the *WISP-1* locus.

We next examined whether the *WISP* genes were amplified in a panel of 25 primary human colon adenocarcinomas. The relative *WISP* gene copy number in each colon tumor DNA was compared with pooled normal DNA from 10 donors by quantitative PCR (Fig. 6). The copy number of *WISP-1* and *WISP-2* was significantly greater than one, approximately 2-fold for *WISP-1* in about 60% of the tumors and 2- to 4-fold for *WISP-2* in 92% of the tumors ($P < 0.001$ for each). The copy number for *WISP-3* was indistinguishable from one ($P = 0.166$). In addition, the copy number of *WISP-2* was significantly higher than that of *WISP-1* ($P < 0.001$).

The levels of *WISP* transcripts in RNA isolated from 19 adenocarcinomas and their matched normal mucosa were

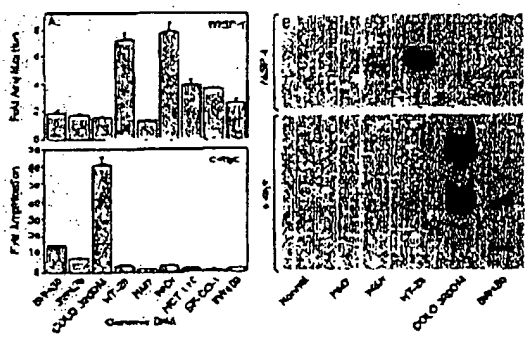


FIG. 5. Amplification of *WISP-1* genomic DNA in colon cancer cell lines (*A*) Amplification in cell line DNA was determined by quantitative PCR. (*B*) Southern blots containing genomic DNA (10 μg) digested with *Eco*RI (*WISP-1*) or *Xba*I (*c-myc*) were hybridized with a 100-bp human *WISP-1* probe (amino acids 186–219) or a human *c-myc* probe (located at bp 1901–2000). The *WISP* and *myc* genes are detected in normal human genomic DNA after a longer film exposure.

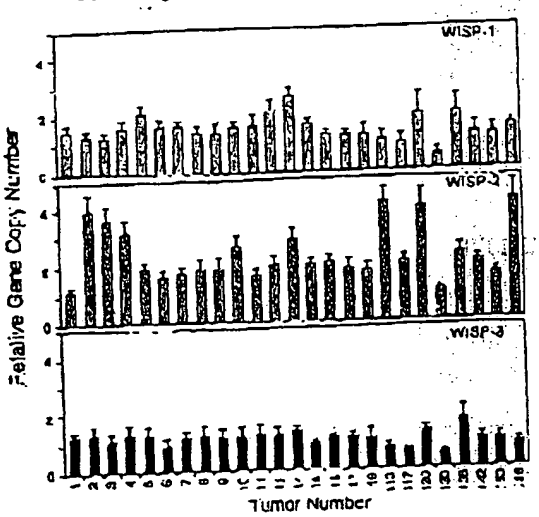


FIG. 6. Genomic amplification of *WISP* genes in human colon tumors. The relative gene copy number of the *WISP* genes in 25 adenocarcinomas was assayed by quantitative PCR, by comparing DNA from primary human tumors with pooled DNA from 10 healthy donors. The data are means ± SEM from one experiment done in triplicate. The experiment was repeated at least three times.

assessed by quantitative PCR (Fig. 7). The level of *WISP-1* RNA present in tumor tissue varied but was significantly increased (2- to >25-fold) in 84% (16/19) of the human colon tumors examined compared with normal adjacent mucosa. Four of 19 tumors showed greater than 10-fold overexpression. In contrast, in 79% (15/19) of the tumors examined, *WISP-2* RNA expression was significantly lower in the tumor than the mucosa. Similar to *WISP-1*, *WISP-3* RNA was overexpressed in 63% (12/19) of the colon tumors compared with the normal

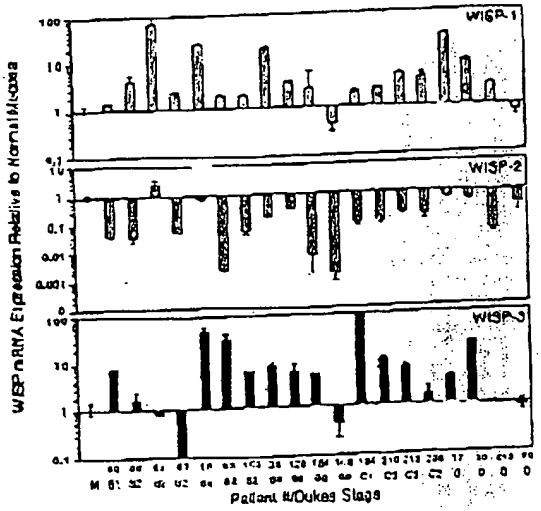


FIG. 7. *WISP* RNA expression in primary human colon tumors relative to expression in normal mucosa from the same patient. Expression of *WISP* mRNA in 19 adenocarcinomas was assayed by quantitative PCR. The Dukes stage of the tumor is listed under the sample number. The data are means ± SEM from one experiment done in triplicate. The experiment was repeated at least twice.

mucosa. The amount of overexpression of *WISP-3* ranged from 4- to >40-fold.

DISCUSSION

One approach to understanding the molecular basis of cancer is to identify differences in gene expression between cancer cells and normal cells. Strategies based on assumptions that steady-state mRNA levels will differ between normal and malignant cells have been used to clone differentially expressed genes (31). We have used a PCR-based selection strategy, SSH, to identify genes selectively expressed in C57MG mouse mammary epithelial cells transformed by Wnt-1.

Three of the genes isolated, *WISP-1*, *WISP-2*, and *WISP-3*, are members of the CCN family of growth factors, which includes CTGF, Cyr61, and *nov*, a family not previously linked to Wnt signaling.

Two independent experimental systems demonstrated that *WISP* induction was associated with the expression of Wnt-1. The first was C57MG cells infected with a Wnt-1 retroviral vector or C57MG cells expressing Wnt-1 under the control of a tetracycline-repressible promoter, and the second was in Wnt-1 transgenic mice, where breast tissue expresses Wnt-1, whereas normal breast tissue does not. No *WISP* RNA expression was detected in mammary tumors induced by polyoma virus middle T antigen (data not shown). These data suggest a link between Wnt-1 and *WISPs* in that in these two situations, *WISP* induction was correlated with Wnt-1 expression.

It is not clear whether the *WISPs* are directly or indirectly induced by the downstream components of the Wnt-1 signaling pathway (i.e., β-catenin-TCF-1/Lef1). The increased levels of *WISP* RNA were measured in Wnt-1-transformed cells, hours or days after Wnt-1 transformation. Thus, *WISP* expression could result from Wnt-1 signaling directly through β-catenin transcription factor regulation or alternatively through Wnt-1 signaling turning on a transcription factor, which in turn regulates *WISPs*.

The *WISPs* define an additional subfamily of the CCN family of growth factors. One striking difference observed in the protein sequence of WISP-2 is the absence of a CT domain, which is present in CTGF, Cyr61, *nov*, WISP-1, and WISP-3. This domain is thought to be involved in receptor binding and dimerization. Growth factors, such as TGF-β, platelet-derived growth factor, and nerve growth factor, which contain a cystine knot motif exist as dimers (32). It is tempting to speculate that WISP-1 and WISP-3 may exist as dimers, whereas WISP-2 exists as a monomer. If the CT domain is also important for receptor binding, WISP-2 may bind its receptor through a different region of the molecule than the other CCN family members. No specific receptors have been identified for CTGF or *nov*. A recent report has shown that integrin $\alpha v\beta_3$ serves as an adhesion receptor for Cyr61 (33).

The strong expression of *WISP-1* and *WISP-2* in cells lying within the fibrovascular tumor stroma in breast tumors from Wnt-1 transgenic animals is consistent with previous observations that transcripts for the related CTGF gene are primarily expressed in the fibrous stroma of mammary tumors (34). Epithelial cells are thought to control the proliferation of connective tissue stroma in mammary tumors by a cascade of growth factor signals similar to that controlling connective tissue formation during wound repair. It has been proposed that mammary tumor cells or inflammatory cells at the tumor interstitial interface secrete TGF-β1, which is the stimulus for stromal proliferation (34). TGF-β1 is secreted by a large percentage of malignant breast tumors and may be one of the growth factors that stimulates the production of CTGF and WISPs in the stroma.

It was of interest that *WISP-1* and *WISP-2* expression was observed in the stromal cells that surrounded the tumor cells

(epithelial cells) in the Wnt-1 transgenic mouse sections of breast tissue. This finding suggests that paracrine signaling could occur in which the stromal cells could supply WISP-1 and WISP-2 to regulate tumor cell growth on the WISP extracellular matrix. Stromal cell-derived factors in the extracellular matrix have been postulated to play a role in tumor cell migration and proliferation (35). The localization of *WISP-1* and *WISP-2* in the stromal cells of breast tumors supports this paracrine model.

An analysis of *WISP-1* gene amplification and expression in human colon tumors showed a correlation between DNA amplification and overexpression, whereas overexpression of *WISP-3* RNA was seen in the absence of DNA amplification. In contrast, *WISP-2* DNA was amplified in the colon tumors, but its mRNA expression was significantly reduced in the majority of tumors compared with the expression in normal colonic mucosa from the same patient. The gene for human *WISP-2* was localized to chromosome 20q12–20q13, at a region frequently amplified and associated with poor prognosis in node negative breast cancer and many colon cancers, suggesting the existence of one or more oncogenes at this locus (36–38). Because the center of the 20q13 amplicon has not yet been identified, it is possible that the apparent amplification observed for *WISP-2* may be caused by another gene in this amplicon.

A recent manuscript on *rCop-1*, the rat orthologue of *WISP-2*, describes the loss of expression of this gene after cell transformation, suggesting it may be a negative regulator of growth in cell lines (16). Although the mechanism by which *WISP-2* RNA expression is down-regulated during malignant transformation is unknown, the reduced expression of *WISP-2* in colon tumors and cell lines suggests that it may function as a tumor suppressor. These results show that the *WISP* genes are aberrantly expressed in colon cancer and suggest that their altered expression may confer selective growth advantage to the tumor.

Members of the Wnt signaling pathway have been implicated in the pathogenesis of colon cancer, breast cancer, and melanoma, including the tumor suppressor gene adenomatous polyposis coli and β-catenin (39). Mutations in specific regions of either gene can cause the stabilization and accumulation of cytoplasmic β-catenin, which presumably contributes to human carcinogenesis through the activation of target genes such as the *WISPs*. Although the mechanism by which Wnt-1 transforms cells and induces tumorigenesis is unknown, the identification of *WISPs* as genes that may be regulated downstream of Wnt-1 in C57MG cells suggests they could be important mediators of Wnt-1 transformation. The amplification and altered expression patterns of the *WISPs* in human colon tumors may indicate an important role for these genes in tumor development.

We thank the DNA synthesis group for oligonucleotide synthesis, T. Baker for technical assistance, F. Dowd for radiation hybrid mapping, K. Willert and R. Nusse for the tet-repressible C57MG/Wnt-1 cells, V. Dixit for discussions, and D. Wood and A. Bruce for artwork.

1. Cadigan, K. M. & Nusse, R. (1997) *Genes Dev.* 11, 3286–3305.
2. Dale, T. C. (1998) *Biochem. J.* 329, 209–223.
3. Nusse, R. & Varmus, H. E. (1982) *Cell* 31, 99–109.
4. van Ooyen, A. & Nusse, R. (1984) *Cell* 39, 233–240.
5. Tsukamoto, A. S., Grosschedl, R., Guzman, R. C., Parslow, T. & Varmus, H. E. (1988) *Cell* 55, 619–625.
6. Brown, J. D. & Moon, R. T. (1998) *Curr. Opin. Cell. Biol.* 10, 182–187.
7. Molenaar, M., van de Wetering, M., Oosterwegel, M., Peterson-Maduro, J., Godsave, S., Korinek, V., Roose, J., Destree, O. & Clevers, H. (1996) *Cell* 86, 391–399.
8. Korinek, V., Barker, N., Willert, K., Molenaar, M., Roose, J., Wagenaar, G., Markman, M., Lamers, W., Destree, O. & Clevers, H. (1998) *Mol. Cell. Biol.* 18, 1248–1256.
9. Matsumine, S., Albert, I., Souza, B., Rubinfeld, B. & Polakis, P. (1995) *Proc. Natl. Acad. Sci. USA* 92, 3030–3034.
10. He, T. C., Sparks, A. B., Rago, C., Hermeking, H., Zawel, L., da Costa, L. T., Morin, P. J., Vogelstein, B. & Kinzler, K. W. (1998) *Science* 281, 1509–1512.
11. Diatchenko, L., Lau, Y. F., Campbell, A. P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E. D. & Siebert, P. D. (1996) *Proc. Natl. Acad. Sci. USA* 93, 6025–6030.
12. Brown, A. M., Wildin, R. S., Prendergast, T. J. & Varmus, H. E. (1986) *Cell* 46, 1001–1009.
13. Wong, G. T., Gavin, B. J. & McMahon, A. P. (1994) *Mol. Cell. Biol.* 14, 6278–6286.
14. Shimizu, H., Julius, M. A., Giarre, M., Zheng, Z., Brown, A. M. & Kitajewski, J. (1997) *Cell Growth Differ.* 8, 1349–1358.
15. Hashimoto, Y., Shindo-Okada, N., Tani, M., Nagamachi, Y., Takeuchi, K., Shiroishi, T., Toma, H. & Yukota, J. (1998) *J. Exp. Med.* 187, 289–296.
16. Zhang, R., Averboukh, L., Zhu, W., Zhang, H., Jo, H., Dempsey, P. J., Coffey, R. J., Pardee, A. B. & Liang, P. (1998) *Mol. Cell. Biol.* 18, 6131–6141.
17. Grotendorst, G. R. (1997) *Cytokine Growth Factor Rev.* 8, 171–179.
18. Kireeva, M. L., Mo, F. E., Yang, G. P. & Lau, L. F. (1996) *Mol Cell. Biol.* 16, 1326–1334.
19. Babic, A. M., Kireeva, M. L., Kolesnikova, T. V. & Lau, L. F. (1998) *Proc. Natl. Acad. Sci. USA* 95, 6355–6360.
20. Martinerie, C., Huff, V., Joubert, I., Haqqioui, M., Saunders, G., Strong, L. & Perbal, B. (1994) *Oncogene* 9, 2729–2732.
21. Bork, P. (1993) *FEBS Lett.* 327, 125–130.
22. Kim, H. S., Nagalla, S. R., Oh, Y., Wilson, E., Roberts, C. T., Jr. & Rosenfeld, R. G. (1997) *Proc. Natl. Acad. Sci. USA* 94, 12981–12986.
23. Joliot, V., Martinerie, C., Dambrine, G., Plassiart, G., Brisac, M., Crochet, J. & Perbal, B. (1992) *Mol. Cell. Biol.* 12, 10–21.
24. Mancuso, D. J., Tuley, E. A., Westfield, L. A., Worrall, N. K., Shelton-Inloes, B. B., Sorace, J. M., Alevy, Y. G. & Sadler, J. E. (1989) *J. Biol. Chem.* 264, 19514–19527.
25. Holt, O. D., Pazeburn, M. K. & Ginsburg, V. (1990) *J. Biol. Chem.* 265, 2852–2855.
26. Voorberg, J., Fontijn, R., Calafat, J., Janssen, H., van Mourik, J. A. & Pannekoek, H. (1991) *J. Cell. Biol.* 113, 195–205.
27. Martinerie, C., Viegas-Pequignot, E., Guenard, I., Dutrillaux, B., Nguyen, V. C., Bernheim, A. & Perbal, B. (1992) *Oncogene* 7, 2529–2534.
28. Takahashi, E., Hori, T., O'Connell, P., Leppert, M. & White, R. (1991) *Cytogenet. Cell. Genet.* 57, 109–111.
29. Mosso, E., Meltzer, P. S., Witkowski, C. M. & Trent, J. M. (1989) *Genes Chromosomes Cancer* 1, 88–94.
30. Carte, S. J. (1994) *Crit. Rev. Oncog.* 4, 435–449.
31. Zhang, L., Zhou, W., Velculescu, V. E., Kern, S. E., Hruban, R. H., Hamilton, S. R., Vogelstein, B. & Kinzler, K. W. (1997) *Science* 276, 1268–1272.
32. Sun, P. D. & Davies, D. R. (1995) *Annu. Rev. Biophys. Biomol. Struct.* 24, 269–291.
33. Kireeva, M. L., Lam, S. C. T. & Lau, L. F. (1998) *J. Biol. Chem.* 273, 3050–3056.
34. Frazier, K. S. & Grotendorst, G. R. (1997) *Int. J. Biochem. Cell Biol.* 29, 153–161.
35. Werner, N. (1997) *Virchows Arch.* 430, 433–441.
36. Tanner, M. M., Tirkkonen, M., Kallioniemi, A., Collins, C., Stokke, T., Karhu, R., Kowbel, D., Shadravan, F., Hintz, M., Kuo, W. L., *et al.* (1994) *Cancer Res.* 54, 4257–4260.
37. Brinkmann, U., Gallo, M., Polymeropoulos, M. H. & Pastan, I. (1996) *Genome Res.* 6, 187–194.
38. Bieche, I., Anderson, L., Zhu, Y., Mossie, K., Ng, L., Souza, B., Schryver, B., Flanagan, P., Clairvoyant, F., Ginther, C., *et al.* (1998) *EMBO J.* 17, 3052–3065.
39. Morin, P. J., Sparks, A. B., Korinek, V., Barker, N., Clevers, H., Vogelstein, B. & Kinzler, K. W. (1997) *Science* 275, 1787–1790.
40. Lu, L. H. & Gillett, N. (1994) *Cell Vision* 1, 169–176.

HellerEhrman
ATTORNEYS

GENOME METHODS

Real Time Quantitative PCR

Christian A. Heid,[1] Junko Stevens,[2] Kenneth J. Livak,[2] and P. Mickey Williams[1,3]

[1]BioAnalytical Technology Department, Genentech, Inc., South San Francisco, California 94080; [2]Applied BioSystems Division of Perkin Elmer Corp., Foster City, California 94404

We have developed a novel "real time" quantitative PCR method. The method measures PCR product accumulation through a dual-labeled fluorogenic probe (i.e., TaqMan Probe). This method provides very accurate and reproducible quantitation of gene copies. Unlike other quantitative PCR methods, real-time PCR does not require post-PCR sample handling, preventing potential PCR product carry-over contamination and resulting in much faster and higher throughput assays. The real-time PCR method has a very large dynamic range of starting target molecule determination (at least five orders of magnitude). Real-time quantitative PCR is extremely accurate and less labor-intensive than current quantitative PCR methods.

Quantitative nucleic acid sequence analysis has had an important role in many fields of biological research. Measurement of gene expression (RNA) has been used extensively in monitoring biological responses to various stimuli (Pan et al. 1994; Huang et al. 1995a,b; Prud'homme et al. 1995). Quantitative gene analysis (DNA) has been used to determine the genomic quantity of a particular gene, as in the case of the human *HER2* gene, which is amplified in ~30% of breast tumors (Slamon et al. 1987). Gene and genome quantitation (DNA and RNA) also have been used for analysis of human immunodeficiency virus (HIV) burden demonstrating changes in the levels of virus throughout the different phases of the disease (Connor et al. 1993; Piatak et al. 1993b; Furtado et al. 1995).

Many methods have been described for the quantitative analysis of nucleic acid sequences (both for RNA and DNA; Southern 1975; Sharp et al. 1980; Thomas 1980). Recently, PCR has proven to be a powerful tool for quantitative nucleic acid analysis. PCR and reverse transcriptase (RT)–PCR have permitted the analysis of minimal starting quantities of nucleic acid (as little as one cell equivalent). This has made possible many experiments that could not have been performed with traditional methods. Although PCR has provided a powerful tool, it is imperative that it be used properly for quantitation (Raymaekers 1995). Many early reports of quantitative PCR and RT–PCR described quantitation of the PCR product but did not measure the initial target sequence quantity. It is essential to design proper controls for the quantitation of the initial target sequences (Ferre 1992; Clementi et al. 1993).

Researchers have developed several methods of quantitative PCR and RT–PCR. One approach measures PCR product quantity in the log phase of the reaction before the plateau (Kellogg et al. 1990; Pang et al. 1990). This method requires that each sample has equal input amounts of nucleic acid and that each sample under analysis amplifies with identical efficiency up to the point of quantitative analysis. A gene sequence (contained in all samples at relatively constant quantities, such as β-actin) can be used for sample amplification efficiency normalization. Using conventional methods of PCR detection and quantitation (gel electrophoresis or plate capture hybridization), it is extremely laborious to assure that all samples are analyzed during the log phase of the reaction (for both the target gene and the normalization gene). Another method, quantitative competitive (QC)–PCR, has been developed and is used widely for PCR quantitation. QC–PCR relies on the inclusion of an internal control competitor in each reaction (Barker-Andre 1991; Piatak et al. 1993a,b). The efficiency of each reaction is normalized to the internal competitor. A known amount of internal competitor can be

[3]Corresponding author.

added to each sample. To obtain relative quantitation, the unknown target PCR product is compared with the known competitor PCR product. Success of a quantitative competitive PCR assay relies on developing an internal control that amplifies with the same efficiency as the target molecule. The design of the competitor and the validation of amplification efficiencies require a dedicated effort. However, because QC–PCR does not require that PCR products be analyzed during the log phase of the amplification, it is the easier of the two methods to use.

Several detection systems are used for quantitative PCR and RT–PCR analysis: (1) agarose gels, (2) fluorescent labeling of PCR products and detection with laser-induced fluorescence using capillary electrophoresis (Fuoco et al. 1995; Williams et al. 1996) or acrylamide gels, and (3) plate capture and sandwich probe hybridization (Mulder et al. 1994). Although these methods proved successful, each method requires post-PCR manipulations that add time to the analysis and may lead to laboratory contamination. The sample throughput of these methods is limited (with the exception of the plate capture approach), and, therefore, these methods are not well suited for uses demanding high sample throughput (i.e., screening of large numbers of biomolecules or analyzing samples for diagnostics or clinical trials).

Here we report the development of a novel assay for quantitative DNA analysis. The assay is based on the use of the 5′ nuclease assay first described by Holland et al. (1991). The method uses the 5′ nuclease activity of *Taq* polymerase to cleave a nonextendible hybridization probe during the extension phase of PCR. The approach uses dual-labeled fluorogenic hybridization probes (Lee et al. 1993; Bassler et al. 1995; Livak et al. 1995a,b). One fluorescent dye serves as a reporter (FAM (i.e., 6-carboxyfluorescein)) and its emission spectra is quenched by the second fluorescent dye, TAMRA (i.e., 6-carboxy-tetramethylrhodamine). The nuclease degradation of the hybridization probe releases the quenching of the FAM fluorescent emission, resulting in an increase in peak fluorescent emission at 518 nm. The use of a sequence detector (ABI Prism) allows measurement of fluorescent spectra of all 96 wells of the thermal cycler continuously during the PCR amplification. Therefore, the reactions are monitored in real time. The output data is described and quantitative analysis of input target DNA sequences is discussed below.

RESULTS

PCR Product Detection in Real Time

The goal was to develop a high-throughput, sensitive, and accurate gene quantitation assay for use in monitoring lipid mediated therapeutic gene delivery. A plasmid encoding human factor VIII gene sequence, pIBTM (see Methods), was used as a model therapeutic gene. The assay uses fluorescent Taqman methodology and an instrument capable of measuring fluorescence in real time (ABI Prism 7700 Sequence Detector). The Taqman reaction requires a hybridization probe labeled with two different fluorescent dyes. One dye is a reporter dye (FAM), the other is a quenching dye (TAMRA). When the probe is intact, fluorescent energy transfer occurs and the reporter dye fluorescent emission is absorbed by the quenching dye (TAMRA). During the extension phase of the PCR cycle, the fluorescent hybridization probe is cleaved by the 5′–3′ nucleolytic activity of the DNA polymerase. On cleavage of the probe, the reporter dye emission is no longer transferred efficiently to the quenching dye, resulting in an increase of the reporter dye fluorescent emission spectra. PCR primers and probes were designed for the human factor VIII sequence and human β-actin gene (as described in Methods). Optimization reactions were performed to choose the appropriate probe and magnesium concentrations yielding the highest intensity of reporter fluorescent signal without sacrificing specificity. The instrument uses a charge-coupled device (i.e., CCD camera) for measuring the fluorescent emission spectra from 500 to 650 nm. Each PCR tube was monitored sequentially for 25 msec with continuous monitoring throughout the amplification. Each tube was re-examined every 8.5 sec. Computer software was designed to examine the fluorescent intensity of both the reporter dye (FAM) and the quenching dye (TAMRA). The fluorescent intensity of the quenching dye, TAMRA, changes very little over the course of the PCR amplification (data not shown). Therefore, the intensity of TAMRA dye emission serves as an internal standard with which to normalize the reporter dye (FAM) emission variations. The software calculates a value termed ΔRn (or ΔRQ) using the following equation: $\Delta Rn = (Rn^+) - (Rn^-)$, where Rn^+ = emission intensity of reporter/emission intensity of quencher at any given time in a reaction tube, and Rn^- = emission intensity of re-

porter/emission intensity of quencher measured prior to PCR amplification in that same reaction tube. For the purpose of quantitation, the last three data points (ΔRns) collected during the extension step for each PCR cycle were analyzed. The nucleolytic degradation of the hybridization probe occurs during the extension phase of PCR, and, therefore, reporter fluorescent emission increases during this time. The three data points were averaged for each PCR cycle and the mean value for each was plotted in an "amplification plot" shown in Figure 1A. The ΔRn mean value is plotted on the *y*-axis, and time, represented by cycle number, is plotted on the *x*-axis. During the early cycles of the PCR amplification, the ΔRn value remains at base line. When sufficient hybridization probe has been cleaved by the *Taq* polymerase nuclease activity, the intensity of reporter fluorescent emission increases. Most PCR amplifications reach a plateau phase of reporter fluorescent emission if the reaction is carried out to high cycle numbers. The amplification plot is examined early in the reaction, at a point that represents the log phase of product accumulation. This is done by assigning an arbitrary threshold that is based on the variability of the base-line data. In Figure 1A, the threshold was set at 10 standard deviations above the mean of base line emission calculated from cycles 1 to 15. Once the threshold is chosen, the point at which

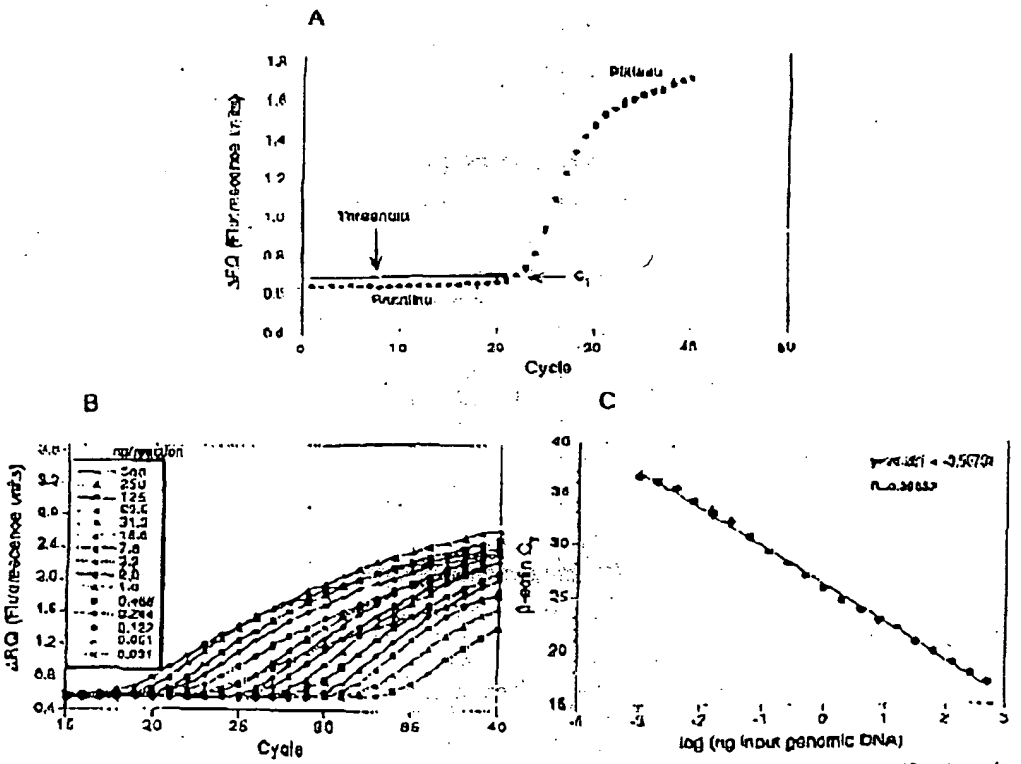


Figure 1 PCR product detection in real time. (*A*) The Model 7700 software will construct amplification plots from the extension phase fluorescent emission data collected during the PCR amplification. The standard deviation is determined from the data points collected from the base line of the amplification plot. C_T values are calculated by determining the point at which the fluorescence exceeds a threshold limit (usually 10 times the standard deviation of the base line). (*B*) Overlay of amplification plots of serially (1:2) diluted human genomic DNA samples amplified with β-actin primers. (*C*) Input DNA concentration of the samples plotted versus C_T. All

the amplification plot crosses the threshold is defined as C_T. C_T is reported as the cycle number at this point. As will be demonstrated, the C_T value is predictive of the quantity of input target.

C_T Values Provide a Quantitative Measurement of Input Target Sequences

Figure 1B shows amplification plots of 15 different PCR amplifications overlaid. The amplifications were performed on a 1:2 serial dilution of human genomic DNA. The amplified target was human β actin. The amplification plots shift to the right (to higher threshold cycles) as the input target quantity is reduced. This is expected because reactions with fewer starting copies of the target molecule require greater amplification to degrade enough probe to attain the threshold fluorescence. An arbitrary threshold of 10 standard deviations above the base line was used to determine the C_T values. Figure 1C represents the C_T values plotted versus the sample dilution value. Each dilution was amplified in triplicate PCR amplifications and plotted as mean values with error bars representing one standard deviation. The C_T values decrease linearly with increasing target quantity. Thus, C_T values can be used as a quantitative measurement of the input target number. It should be noted that the amplification plot for the 15.6-ng sample shown in Figure 1B does not reflect the same fluorescent rate of increase exhibited by most of the other samples. The 15.6-ng sample also achieves endpoint plateau at a lower fluorescent value than would be expected based on the input DNA. This phenomenon has been observed occasionally with other samples (data not shown) and may be attributable to late cycle inhibition; this hypothesis is still under investigation. It is important to note that the flattened slope and early plateau do not impact significantly the calculated C_T value as demonstrated by the fit on the line shown in Figure 1C. All triplicate amplifications resulted in very similar C_T values—the standard deviation did not exceed 0.5 for any dilution. This experiment contains a >100,000-fold range of input target molecules. Using C_T values for quantitation permits a much larger assay range than directly using total fluorescent emission intensity for quantitation. The linear range of fluorescent intensity measurement of the ABI Prism 7700 Sequence Detector begins to fall off at high fluorescence values, whereas C_T values permit quantitative measurements over a very large range of relative starting target quantities.

Sample Preparation Validation

Several parameters influence the efficiency of PCR amplification: magnesium and salt concentrations, reaction conditions (i.e., time and temperature), PCR target size and composition, primer sequences, and sample purity. All of the above factors are common to a single PCR assay, except sample to sample purity. In an effort to validate the method of sample preparation for the factor VIII assay, PCR amplification reproducibility and efficiency of 10 replicate sample preparations were examined. After genomic DNA was prepared from the 10 replicate samples, the DNA was quantitated by ultraviolet spectroscopy. Amplifications were performed analyzing β-actin gene content in 100 and 25 ng of total genomic DNA. Each PCR amplification was performed in triplicate. Comparison of C_T values for each triplicate sample show minimal variation based on standard deviation and coefficient of variance (Table 1). Therefore, each of the triplicate PCR amplifications was highly reproducible, demonstrating that real time PCR using this instrumentation introduces minimal variation into the quantitative PCR analysis. Comparison of the mean C_T values of the 10 replicate sample preparations also showed minimal variability, indicating that each sample preparation yielded similar results for β-actin gene quantity. The highest C_T difference between any of the samples was 0.85 and 0.71 for the 100 and 25 ng samples, respectively. Additionally, the amplification of each sample exhibited an equivalent rate of fluorescent emission intensity change per amount of DNA target analyzed as indicated by similar slopes derived from the sample dilutions (Fig. 2). Any sample containing an excess of a PCR inhibitor would exhibit a greater measured β-actin C_T value for a given quantity of DNA. In addition, the inhibitor would be diluted along with the sample in the dilution analysis (Fig. 2), altering the expected C_T value change. Each sample amplification yielded a similar result in the analysis, demonstrating that this method of sample preparation is highly reproducible with regard to sample purity.

Quantitative Analysis of a Plasmid After

Table 1. Reproducibility of Sample Preparation Method

	100 ng				25 ng			
Sample no.	C_T	mean	standard deviation	CV	C_T	mean	standard deviation	CV
1	18.24				20.48			
	18.23				20.55			
	18.33	18.27	0.06	0.32	20.5	20.51	0.03	0.17
2	18.33				20.61			
	18.35				20.59			
	18.44	18.37	0.06	0.32	20.41	20.54	0.11	0.54
3	18.3				20.54			
	18.3				20.6			
	18.42	18.34	0.07	0.36	20.49	20.54	0.06	0.28
4	18.15				20.48			
	18.23				20.44			
	18.32	18.23	0.08	0.46	20.38	20.43	0.05	0.26
5	18.4				20.68			
	18.38				20.87			
	18.46	18.42	0.04	0.23	20.63	20.71	0.13	0.61
6	18.54				21.09			
	18.67				21.04			
	19	18.74	0.24	1.26	21.01	21.06	0.03	0.15
7	18.28				20.67			
	18.36				20.73			
	18.52	18.39	0.12	0.66	20.65	20.68	0.04	0.2
8	18.45				20.98			
	18.7				20.84			
	18.73	18.63	0.16	0.83	20.75	20.86	0.12	0.57
9	18.18				20.46			
	18.36				20.54			
	18.36	18.29	0.1	0.55	20.48	20.51	0.07	0.32
10	18.42				20.79			
	18.57				20.78			
	18.66	18.55	0.12	0.66	20.62	20.73	0.1	0.46
Mean	(1-10)	18.42	0.17	0.90		20.66	0.19	0.94

for containing a partial cDNA for human factor VIII, pHGFM. A series of transfections was set up using a decreasing amount of the plasmid (40, 4, 0.5, and 0.1 μg). Twenty-four hours post-transfection, total DNA was purified from each flask of cells. β-Actin gene quantity was chosen as a value for normalization of genomic DNA concentration from each sample. In this experiment, β-actin gene content should remain constant relative to total genomic DNA. Figure 3 shows the result of the β-actin DNA measurement (100 ng total DNA determined by ultraviolet spectroscopy) of each sample. Each sample was analyzed in triplicate and the mean β-actin C_T values of the triplicates were plotted (error bars represent one standard deviation). The highest difference between any two sample means was 0.95 C_T. Ten nanograms of total DNA of each sample were also examined for β-actin. The results again showed that very similar amounts of genomic DNA were present; the maximum mean β-actin C_T value difference was 1.0. As Figure 3 shows, the rate of β-actin C_T change between the 100- and 10-ng samples was similar (slope values range between −3.56 and −3.45). This verifies again that the method of sample preparation yields samples of identical PCR integrity (i.e., no sample contained an excessive amount of a PCR inhibitor). However, these results indicate that each sample contained slight differences in the actual amount of genomic DNA analyzed. Determination of actual genomic DNA concentration was accomplished

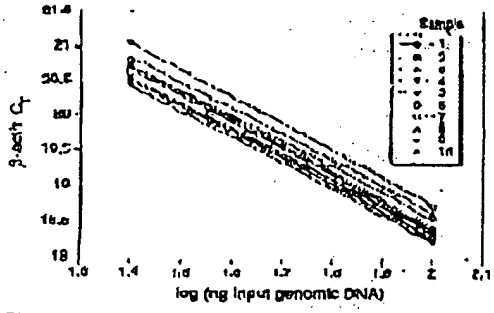


Figure 2 Sample preparation purity. The replicate samples shown in Table 1 were also amplified in triplicate using 25 ng of each DNA sample. The figure shows the input DNA concentration (100 and 25 ng) vs. C_T. In the figure, the 100 and 25 ng points for each sample are connected by a line.

by plotting the mean β-actin C_T value obtained for each 100-ng sample on a β-actin standard curve (shown in Fig. 4C). The actual genomic DNA concentration of each sample, *a*, was obtained by extrapolation to the *x*-axis.

Figure 4A shows the measured (i.e., non-normalized) quantities of factor VIII plasmid DNA (pF8TM) from each of the four transient cell transfections. Each reaction contained 100 ng of total sample DNA (as determined by UV spectroscopy). Each sample was analyzed in triplicate

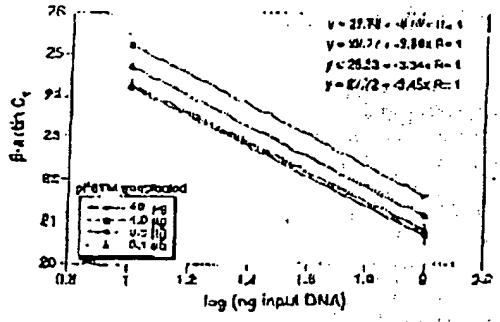


Figure 3 Analysis of transfected cell DNA quantity and purity. The DNA preparations of the four 293 cell transfections (40, 4, 0.5, and 0.1 μg of pF8TM) were analyzed for the β-actin gene, 100 and 10 ng (determined by ultraviolet spectroscopy) of each sample were amplified in triplicate. For each amount of pF8TM that was transfected, the β-actin C_T values are plotted versus the total input DNA concentration.

PCR amplifications. As shown, pF8TM purified from the 293 cells decreases (mean C_T values increased) with decreasing amounts of plasmid transfected. The mean C_T values obtained for pF8TM in Figure 4A were plotted on a standard curve comprised of serially diluted pF8TM, shown in Figure 4B. The quantity of pF8TM, *a*, found in each of the four transfections was determined by extrapolation to the *x* axis of the standard curve in Figure 4B. These uncorrected values, *b*, for pF8TM were normalized to determine the actual amount of pF8TM found per 100 ng of genomic DNA by using the equation:

$$\frac{b \times 100\ \text{ng}}{a} = \begin{array}{l}\text{actual pF8TM copies per}\\ \text{100 ng of genomic DNA}\end{array}$$

where *a* = actual genomic DNA in a sample and *b* = pF8TM copies from the standard curve. The normalized quantity of pF8TM per 100 ng of genomic DNA for each of the four transfections is shown in Figure 4D. These results show that the quantity of factor VIII plasmid associated with the 293 cells, 24 hr after transfection, decreases with decreasing plasmid concentration used in the transfection. The quantity of pF8TM associated with 293 cells, after transfection with 40 μg of plasmid, was 35 pg per 100 ng genomic DNA. This results in ~520 plasmid copies per cell.

DISCUSSION

We have described a new method for quantitating gene copy numbers using real-time analysis of PCR amplifications. Real-time PCR is compatible with either of the two PCR (RT–PCR) approaches (1) quantitative competitive where an internal competitor for each target sequence is used for normalization (data not shown) or (2) quantitative comparative PCR using a normalization gene contained within the sample (i.e., β-actin) or a "housekeeping" gene for RT–PCR. If equal amounts of nucleic acid are analyzed for each sample and if the amplification efficiency before quantitative analysis is identical for each sample, the internal control (normalization gene or competitor) should give equal signals for all samples.

The real-time PCR method offers several advantages over the other two methods currently employed (see the Introduction). First, the real-time PCR method is performed in a closed-tube system and requires no post-PCR manipulation

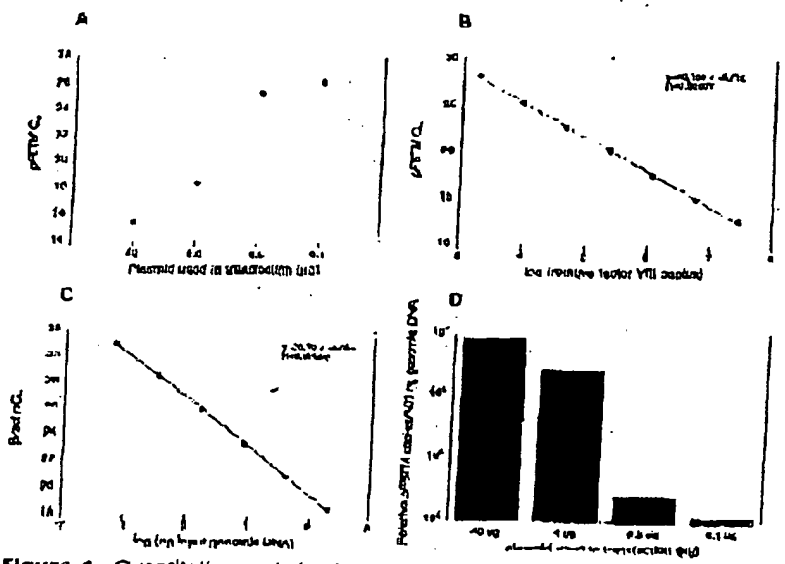


Figure 4 Quantitative analysis of pFBTM in transfected cells. (*A*) Amount of plasmid DNA used for the transfection plotted against the mean C_T value determined for pFBTM remaining 24 hr after transfection. (*B*,*C*) Standard curves of pFATM and β-actin, respectively. pFBTM DNA (*B*) and genomic DNA (*C*) were diluted serially 1:5 before amplification with the appropriate primers. The β-actin standard curve was used to normalize the results of *A* to 100 ng of genomic DNA. (*D*) The amount of pFBTM present per 100 ng of genomic DNA.

of sample. Therefore, the potential for PCR contamination in the laboratory is reduced because amplified products can be analyzed and disposed of without opening the reaction tubes. Second, this method supports the use of a normalization gene (i.e., β-actin) for quantitative PCR or housekeeping genes for quantitative RT–PCR controls. Analysis is performed in real time during the log phase of product accumulation. Analysis during log phase permits many different genes (over a wide input target range) to be analyzed simultaneously, without concern of reaching reaction plateau at different cycles. This will make multigene analysis assays much easier to develop, because individual internal competitors will not be needed for each gene under analysis. Third, sample throughput will increase dramatically with the new method because there is no post-PCR processing time. Additionally, working in a 96-well format is highly compatible with automation technology.

The real-time PCR method is highly reproducible. Replicate amplifications can be analyzed for each sample minimizing potential error. The system allows for a very large assay dynamic range (approaching 1,000,000-fold starting target). Using a standard curve for the target of interest, relative copy number values can be determined for any unknown sample. Fluorescent threshold values, C_T, correlate linearly with relative DNA copy numbers. Real time quantitative RT-PCR methodology (Gibson et al., this issue) has also been developed. Finally, real time quantitative PCR methodology can be used to develop high-throughput screening assays for a variety of applications (quantitative gene expression (RT-PCR), gene copy assays (Her2, HIV, etc.), genotyping (knockout mouse analysis), and Immuno-PCR).

Real-time PCR may also be performed using intercalating dyes (Higuchi et al. 1992) such as ethidium bromide. The fluorogenic probe method offers a major advantage over intercalating dyes—greater specificity (i.e., primer dimers and nonspecific PCR products are not detected).

METHODS

Generation of a Plasmid Containing a Partial cDNA for Human Factor VIII

Total RNA was harvested (RNAzol B from Tel Test, Inc., Friendswood, TX) from cells transfected with a factor VIII expression vector, pCIS2.8(2S1) (Eaton et al. 1986; Gorman et al. 1990). A factor VIII partial cDNA sequence was generated by RT PCR (GeneAmp RT PCR kit (part N8080179, PE Applied Biosystems, Foster City, CA)) using the PCR primers FVIIfor and FVIIrev (primer sequences are shown below). The amplicon was reamplified using modified FVIIfor and FVIIrev primers (appended with *Xmnl*I and *Hin*dIII restriction site sequences at the 5' ends) and cloned into pCI2M-3Z (Promega Corp., Madison, WI). The resulting clone, pF8TM, was used for transient transfection of 293 cells.

Amplification of Target DNA and Detection of Amplicon Factor VIII Plasmid DNA

(pF8TM) was amplified with the primers FVIIfor 5'-CCCCTTCGCAAGAGTAACTTCGTC-3' and FVIIrev 5'-AAACCTCACGCTGCATCTTACTC-3'. The reaction produced a 422-bp PCR product. The forward primer was designed to recognize a unique sequence found in the 5' untranslated region of the parent pCIS2.8(2S1) plasmid and therefore does not recognize and amplify the human factor VIII gene. Primers were chosen with the assistance of the computer program Oligo 4.0 (National Biosciences, Inc., Plymouth, MN). The human β-actin gene was amplified with the primers β-actin forward primer 5'-TCACCCACACTGTGCCCATCTACGA-3' and β-actin reverse primer 5'-CAGCGGAACCGCTCATTGCCAATGG-3'. The reaction produced a 295-bp PCR product.

Amplification reactions (50 μl) contained a DNA sample, 10× PCR Buffer II (5 μl), 200 μM dATP, dCTP, dGTP, and 400 μM dUTP, 4 mM $MgCl_2$, 1.25 Units AmpliTaq DNA polymerase, 0.5 unit AmpErase uracil *N*-glycosylase (UNG), 60 pmole of each factor VIII primer, and 15 pmole of each β actin primer. The reaction also contained one of the following detection probes (100 nM each): FVIIprobe 5'-(FAM)AGCTGTTGCAGCTTCCTTTCTTTCTGT-GCCTT(TAMRA)-p 3' and β-actin probe 5'-(FAM)ATGCCC-X(TAMRA)CCCCCATGCCATG-p-3' where p indicates phosphorylation and X indicates a linker arm nucleotide. Reaction tubes were MicroAmp Optical Tubes (part Number N801 0933, Perkin Elmer) that were frosted (at Perkin Elmer) to prevent light from reflecting. Tube caps were similar to MicroAmp Caps but specially designed to prevent light scattering. All of the PCR consumables were supplied by PE Applied Biosystems (Foster City, CA) except the factor VIII primers, which were synthesized at Genentech, Inc. (South San Francisco, CA). Probes were designed using the Oligo 4.0 software, following guidelines suggested in the Model 7700 Sequence Detector instrument manual. Briefly, probe T_m should be at least 5°C higher than the annealing temperature used during thermal cycling; primers should not form stable duplexes with the probe.

The thermal cycling conditions included 2 min at 50°C and 10 min at 95°C. Thermal cycling proceeded with 40 cycles of 95°C for 15 s and 60°C for 1 min. All reactions were performed in the Model 7700 Sequence Detector (PE Applied Biosystems), which contains a GeneAmp PCR System 9600. Reaction conditions were programmed on a Power Macintosh 7100 (Apple Computer, Santa Clara, CA) linked directly to the Model 7700 Sequence Detector. Analysis of data was also performed on the Macintosh computer. Collection and analysis software was developed at PE Applied Biosystems.

Transfection of Cells with Factor VIII Construct

Four T175 flasks of 293 cells (ATCC CRL 1573), a human fetal kidney suspension cell line, were grown to 80% confluency and transfected pF8TM. Cells were grown in the following media: 50% MAMX #12 without GHT, 50% low glucose Dulbecco's modified Eagle medium (DMEM) without glycine with sodium bicarbonate, 10% fetal bovine serum, 2 mM L-glutamine, and 1% penicillin–streptomycin. The media was changed 20 min before the transfection. pF8TM DNA amounts of 40, 4, 0.5, and 0.1 μg were added to 1.6 ml of a solution containing 0.125 × $CaCl_2$ and 1× HBPBS. The four mixtures were left at room temperature for 10 min and then added dropwise to the cells. The flasks were incubated at 37°C and 5% CO_2 for 24 hr, washed with PBS, and resuspended in PBS. The resuspended cells were divided into aliquots and DNA was extracted immediately using the QIAamp Blood Kit (Qiagen, Chatsworth, CA). DNA was eluted into 200 μl of 20 mM Tris-HCl at pH 8.0.

ACKNOWLEDGMENTS

We thank Genentech's DNA Synthesis Group for primer synthesis and Genentech's Graphics Group for assistance with the figures.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Bassler, H.A., S.J. Flood, K.J. Livak, J. Marmaro, R. Knorr, and C.A. Batt. 1995. Use of a fluorogenic probe in a PCR-based assay for the detection of Listeria monocytogenes. *Appl. Environ. Microbiol.* 61: 3724–3728.

Becker-Andre, M. 1991. Quantitative evaluation of mRNA levels. *Meth. Mol. Cell. Biol.* 2: 189–201.

Clementi, M., S. Menzo, P. Bagnarelli, A. Manzin, A. Valenza, and P.E. Varaldo. 1993. Quantitative PCR and RT-PCR in virology. [Review]. *PCR Methods Applic.* 2: 191–196.

Connor, R.I., H. Mohri, Y. Cao, and D.D. Ho. 1993. Increased viral burden and cytopathicity correlate temporally with CD4+ T-lymphocyte decline and clinical progression in human immunodeficiency virus type 1-infected individuals. *J. Virol.* 67: 1772–1777.

Eaton, D.L., W.J. Wood, D. Eaton, P.E. Hass, P.

Vehar, and C. Gorman. 1986. Construction and characterization of an active factor VIII variant lacking the central one third of the molecule. *Biochemistry* 25: 8343–8347.

Fasco, M.J., C.P. Treanor, S. Spivack, H.L. Figge, and L.S. Kaminsky. 1995. Quantitative RNA-polymerase chain reaction-DNA analysis by capillary electrophoresis and laser-induced fluorescence. *Anal. Biochem.* 224: 140–147.

Ferre, F. 1992. Quantitative or semi-quantitative PCR: Reality versus myth. *PCR Methods Applic.* 2: 1–9.

Furtado, M.R., L.A. Kingsley, and S.M. Wolinsky. 1995. Changes in the viral mRNA expression pattern correlate with a rapid rate of CD4+ T-cell number decline in human immunodeficiency virus type 1-infected individuals. *J. Virol.* 69: 2092–2100.

Gibson, U.E.M., C.A. Heid, and P.M. Williams. 1996. A novel method for real time quantitative competitive RT-PCR. *Genome Res.* (this issue).

Gorman, C.M., D.R. Gies, and G. McCray. 1990. Transient production of proteins using an adenovirus transformed cell line. *DNA Prot. Engin. Tech.* 2: 3–10.

Higuchi, R., G. Dollinger, P.S. Walsh, and R. Griffith. 1992. Simultaneous amplification and detection of specific DNA sequences. *Biotechnology* 10: 413–417.

Holland, P.M., R.D. Abramson, R. Watson, and D.H. Gelfand. 1991. Detection of specific polymerase chain reaction product by utilizing the 5'→3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proc. Natl. Acad. Sci.* 88: 7276–7280.

Huang, S.K., H.Q. Xiao, T.J. Kleine, G. Paciotti, D.G. Marsh, L.M. Lichtenstein, and M.C. Liu. 1995a. IL-13 expression at the sites of allergen challenge in patients with asthma. *J. Immun.* 155: 2688–2694.

Huang, S.K., M. Yi, E. Palmer, and D.G. Marsh. 1995b. A dominant T cell receptor beta-chain in response to a short ragweed allergen, Amb a 5. *J. Immun.* 154: 6157–6162.

Kellogg, D.E., J.J. Sninsky, and S. Kwok. 1990. Quantitation of HIV-1 proviral DNA relative to cellular DNA by the polymerase chain reaction. *Anal. Biochem.* 189: 202–208.

Lee, L.G., C.R. Connell, and W. Bloch. 1993. Allelic discrimination by nick-translation PCR with fluorogenic probes. *Nucleic Acids Res.* 21: 3761–3766.

Livak, K.J., S.J. Flood, J. Marmaro, W. Giusti, and K. Deetz. 1995a. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Applic.* 4: 357–362.

Livak, K.J., J. Marmaro, and J.A. Todd. 1995b. Towards fully automated genome-wide polymorphism screening [Letter]. *Nature Genet.* 9: 341–342.

Mulder, J., N. McKinney, C. Christopherson, J. Sninsky, L. Greenfield, and S. Kwok. 1994. Rapid and simple PCR assay for quantitation of human immunodeficiency virus type 1 RNA in plasma: Application to acute retroviral infection. *J. Clin. Microbiol.* 32: 292–300.

Pang, S., Y. Koyanagi, S. Miles, C. Wiley, H.V. Vinters, and I.S. Chen. 1990. High levels of unintegrated HIV-1 DNA in brain tissue of AIDS dementia patients. *Nature* 343: 85–89.

Piatak, M.J., K.C. Luk, B. Williams, and J.D. Lifson. 1993a. Quantitative competitive polymerase chain reaction for accurate quantitation of HIV DNA and RNA species. *BioTechniques* 14: 70–81.

Piatak, M.J., M.S. Saag, L.C. Yang, S.J. Clark, J.C. Kappes, K.C. Luk, B.H. Hahn, G.M. Shaw, and J.D. Lifson. 1993b. High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR [see Comments]. *Science* 259: 1749–1754.

Prud'homme, G.J., D.H. Kono, and A.N. Theofilopoulos. 1995. Quantitative polymerase chain reaction analysis reveals marked overexpression of interleukin-1 beta, interleukin-1 and interferon-gamma mRNA in the lymph nodes of lupus-prone mice. *Mol. Immunol.* 32: 495–503.

Raeymaekers, L. 1995. A commentary on the practical applications of competitive PCR. *Genome Res.* 5: 91–94.

Sharp, P.A., A.J. Berk, and S.M. Berget. 1980. Transcription maps of adenovirus. *Methods Enzymol.* 65: 750–768.

Slamon, D.J., G.M. Clark, S.G. Wong, W.J. Levin, A. Ullrich, and W.L. McGuire. 1987. Human breast cancer: Correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 235: 177–182.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98: 503–517.

Tan, X., X. Sun, C.F. Gonzalez, and W. Hsueh. 1994. PAF and TNF increase the precursor of NF-kappa B p50 mRNA in mouse intestine: Quantitative analysis by competitive PCR. *Biochim. Biophys. Acta* 1215: 157–162.

Thomas, P.S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci.* 77: 5201–5205.

Williams, S., C. Schwer, A. Krishnarao, C. Heid, B. Karger, and P.M. Williams. 1996. Quantitative competitive PCR: Analysis of amplified products of the HIV-1 gag gene by capillary electrophoresis with laser induced fluorescence detection. *Anal. Biochem.* (in press).

Received June 3, 1996; accepted in revised form July 29, 1996.

HellerEhrman
ATTORNEYS

methods. Peptides AENK or AEQK were dissolved in water, made isotonic with NaCl and diluted into RPMI growth medium. T-cell-proliferation assays were done essentially as described[20,21]. Briefly, after antigen pulsing (30 μg ml^{-1} TTCF) with tetrapeptides (1–2 mg ml^{-1}), PBMCs or EBV-B cells were washed in PBS and fixed for 45 s in 0.05% glutaraldehyde. Glycine was added to a final concentration of 0.1M and the cells were washed five times in RPMI 1640 medium containing 1% FCS before co-culture with T-cell clones in round-bottom 96-well microtitre plates. After 48 h, the cultures were pulsed with 1 μCi of ^{3}H-thymidine and harvested for scintillation counting 16 h later. Predigestion of native TTCF was done by incubating 200 μg TTCF with 0.25 μg pig kidney legumain in 500 μl 50 mM citrate buffer, pH 5.5, for 1 h at 37 °C.

Glycopeptide digestions. The peptides HIDNEEDI, HIDN(*N*-glucosamine) EEDI and HIDNESDI, which are based on the TTCF sequence, and QQQHLFGSNVTDCSGNFCLFR(KKK), which is based on human transferrin, were obtained by custom synthesis. The three C-terminal lysine residues were added to the natural sequence to aid solubility. The transferrin glycopeptide QQQHLFGSNVTDCSGNFCLFR was prepared by tryptic (Promega) digestion of 5 mg reduced, carboxy-methylated human transferrin followed by concanavalin A chromatography[14]. Glycopeptides corresponding to residues 622–642 and 421–452 were isolated by reverse-phase HPLC and identified by mass spectrometry and N-terminal sequencing. The lyophilized transferrin-derived peptides were redissolved in 50 mM sodium acetate, pH 5.5, 10 mM dithiothreitol, 20% methanol. Digestions were performed for 3 h at 30 °C with 5–50 mU ml^{-1} pig kidney legumain or B-cell AEP. Products were analysed by HPLC or MALDI-TOF mass spectrometry using a matrix of 10 mg ml^{-1} α-cyanocinnamic acid in 50% acetonitrile/0.1% TFA and a PerSeptive Biosystems Elite STR mass spectrometer set to linear or reflector mode. Internal standardization was obtained with a matrix ion of 568.13 mass units.

Received 29 September; accepted 3 November 1998.

1. Chen, J. M. *et al.* Cloning, isolation, and characterization of mammalian legumain, an asparaginyl endopeptidase. *J. Biol. Chem.* 272, 8090–8098 (1997).
2. Kembhavi, A. A., Buttle, D. J., Knight, C. G. & Barrett, A. J. The two cysteine endopeptidases of legume seeds: purification and characterization by use of specific fluoromethylketone inhibitors. *Arch. Biochem. Biophys.* 303, 208–213 (1993).
3. Dalton, J. P., Hola-Jamriska, L. & Bridley, P. J. Asparaginyl endopeptidase activity in adult *Schistosoma mansoni*. *Parasitology* 111, 575–580 (1995).
4. Bennett, K. *et al.* Antigen processing for presentation by class II major histocompatibility complex requires cleavage by cathepsin E. *Eur. J. Immunol.* 22, 1519–1524 (1992).
5. Riese, R. J. *et al.* Essential role for cathepsin S in MHC class II-associated invariant chain processing and peptide loading. *Immunity* 4, 357–366 (1996).
6. Rodriguez, G. M. & Diment, S. Role of cathepsin D in antigen presentation of ovalbumin. *J. Immunol.* 149, 2894–2898 (1992).
7. Hewitt, E. W. *et al.* Natural processing sites for human cathepsin E and cathepsin D in tetanus toxin: implications for T cell epitope generation. *J. Immunol.* 159, 4693–4699 (1997).
8. Watts, C. Capture and processing of exogenous antigens for presentation on MHC molecules. *Annu. Rev. Immunol.* 15, 821–850 (1997).
9. Chapman, H. A. Endosomal proteases and MHC class II function. *Curr. Opin. Immunol.* 10, 93–102 (1998).
10. Pieters, J. & Müller, J. Endosomal proteases and antigen processing. *Trends Biochem. Sci.* 22, 377–382 (1997).
11. Liu, J. & van Halbeek, H. Complete ^{1}H and ^{13}C resonance assignments of a 21-amino acid glycopeptide prepared from human serum transferrin. *Carbohydr. Res.* 296, 1–21 (1996).
12. Pearson, O. T. & Lackaye, R. M. The instructive role of innate immunity in the acquired immune response. *Science* 272, 50–54 (1996).
13. Medzhitov, R. & Janeway, C. A. J. Innate immunity: the virtues of a nonclonal system of recognition. *Cell* 91, 295–298 (1997).
14. Wyatt, R. *et al.* The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* 393, 705–711 (1998).
15. Mortarelli, P. *et al.* N-glycosylation of HIV gp120 may constrain recognition by T lymphocytes. *J. Immunol.* 147, 3128–3132 (1991).
16. Davidson, H. W., West, M. A. & Watts, C. Endocytosis, intracellular trafficking, and processing of membrane IgG and monovalent antigen/membrane IgG complexes in B lymphocytes. *J. Immunol.* 144, 4101–4109 (1990).
17. Barrett, A. J. & Kirschke, H. Cathepsin B, cathepsin H and cathepsin L. *Methods Enzymol.* 80, 535–559 (1981).
18. Makoff, A. J., Oxbrandine, S. P., Smallwood, A. E. & Fairweather, N. F. Expression of tetanus toxin fragment C in *E. coli*: its purification and potential use as a vaccine. *Biotechnology* 7, 1043–1046 (1989).
19. Lane, D. P. & Harlow, E. *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1988).
20. Lanzavecchia, A. Antigen-specific interaction between T and B cells. *Nature* 314, 537–539 (1985).
21. Pond, L. & Watts, C. Characterization of transport of newly assembled, T cell-stimulatory MHC class II-peptide complexes from MHC class II compartments to the cell surface. *J. Immunol.* 159, 543–553 (1997).

Acknowledgements. We thank M. Ferguson for helpful discussions and advice; E. Smythe and L. Grayson for advice and technical assistance; B. Spence, A. Knight and the BTS (Ninewells Hospital) for help with blood monocyte preparation; and our colleagues for many helpful comments on the manuscript. This work was supported by the Wellcome Trust and by an EMBO Long-term fellowship to B. M.

Correspondence and requests for materials should be addressed to C.W. (e-mail: c.watts@dundee.ac.uk).

Genomic amplification of a decoy receptor for Fas ligand in lung and colon cancer

Robert M. Pitti*†, Scot A. Marsters*†, David A. Lawrence*, Margaret Roy*, Frank C. Kischkel*, Patrick Dowd*, Arthur Huang*, Christopher J. Donahue*, Steven W. Sherwood*, Daryl T. Baldwin*, Paul J. Godowski*, William I. Wood*, Austin L. Gurney*, Kenneth J. Hillan*, Robert L. Cohen*, Audrey D. Goddard*, David Botstein‡ & Avi Ashkenazi*

* *Departments of Molecular Oncology, Molecular Biology, and Immunology, Genentech Inc., 1 DNA Way, South San Francisco, California 94080, USA*
‡ *Department of Genetics, Stanford University, Stanford, California 94305, USA*
† *These authors contributed equally to this work*

Fas ligand (FasL) is produced by activated T cells and natural killer cells and it induces apoptosis (programmed cell death) in target cells through the death receptor Fas/Apo1/CD95 (ref. 1). One important role of FasL and Fas is to mediate immune-cytotoxic killing of cells that are potentially harmful to the organism, such as virus-infected or tumour cells[1]. Here we report the discovery of a soluble decoy receptor, termed decoy receptor 3 (DcR3), that binds to FasL and inhibits FasL-induced apoptosis. The DcR3 gene was amplified in about half of 35 primary lung and colon tumours studied, and DcR3 messenger RNA was expressed in malignant tissue. Thus, certain tumours may escape FasL-dependent immune-cytotoxic attack by expressing a decoy receptor that blocks FasL.

By searching expressed sequence tag (EST) databases, we identified a set of related ESTs that showed homology to the tumour necrosis factor (TNF) receptor (TNFR) gene superfamily[2]. Using the overlapping sequence, we isolated a previously unknown full-length complementary DNA from human fetal lung. We named the protein encoded by this cDNA decoy receptor 3 (DcR3). The cDNA encodes a 300-amino-acid polypeptide that resembles members of the TNFR family (Fig. 1a): the amino terminus contains a leader sequence, which is followed by four tandem cysteine-rich domains (CRDs). Like one other TNFR homologue, osteoprotegerin (OPG)[3], DcR3 lacks an apparent transmembrane sequence, which indicates that it may be a secreted, rather than a membrane-associated, molecule. We expressed a recombinant, histidine-tagged form of DcR3 in mammalian cells; DcR3 was secreted into the cell culture medium, and migrated on polyacrylamide gels as a protein of relative molecular mass 35,000 (data not shown). DcR3 shares sequence identity in particular with OPG (31%) and TNFR2 (29%), and has relatively less homology with Fas (17%). All of the cysteines in the four CRDs of DcR3 and OPG are conserved; however, the carboxy-terminal portion of DcR3 is 101 residues shorter.

We analysed expression of DcR3 mRNA in human tissues by northern blotting (Fig. 1b). We detected a predominant 1.2-kilobase transcript in fetal lung, brain, and liver, and in adult spleen, colon and lung. In addition, we observed relatively high DcR3 mRNA expression in the human colon carcinoma cell line SW480.

To investigate potential ligand interactions of DcR3, we generated a recombinant, Fc-tagged DcR3 protein. We tested binding of DcR3–Fc to human 293 cells transfected with individual TNF-family ligands, which are expressed as type 2 transmembrane proteins (these transmembrane proteins have their N termini in the cytosol). DcR3–Fc showed a significant increase in binding to cells transfected with FasL[4] (Fig. 2a), but not to cells transfected with TNF[5], Apo2L/TRAIL[6,7], Apo3L/TWEAK[8,9], or OPGL/TRANCE/

letters to nature

RANKL[10-12] (data not shown). DcR3–Fc immunoprecipitated shed FasL from FasL-transfected 293 cells (Fig. 2b) and purified soluble FasL (Fig. 2c), as did the Fc-tagged ectodomain of Fas but not TNFR1. Gel-filtration chromatography showed that DcR3–Fc and soluble FasL formed a stable complex (Fig. 2d). Equilibrium analysis indicated that DcR3–Fc and Fas–Fc bound to soluble FasL with a comparable affinity ($K_d = 0.8 \pm 0.2$ and 1.1 ± 0.1 nM, respectively; Fig. 2e), and that DcR3–Fc could block nearly all of the binding of soluble FasL to Fas–Fc (Fig. 2e, inset). Thus, DcR3 competes with Fas for binding to FasL.

To determine whether binding of DcR3 inhibits FasL activity, we tested the effect of DcR3–Fc on apoptosis induction by soluble FasL in Jurkat T leukaemia cells, which express Fas (Fig. 3a). DcR3–Fc and Fas–Fc blocked soluble-FasL-induced apoptosis in a similar dose-dependent manner, with half-maximal inhibition at $-0.1\ \mu g\ ml^{-1}$. Time-course analysis showed that the inhibition did not merely delay cell death, but rather persisted for at least 24 hours (Fig. 3b). We also tested the effect of DcR3–Fc on activation-induced cell death (AICD) of mature T lymphocytes, a FasL-dependent process[1]. Consistent with previous results[13], activation of interleukin-2-stimulated CD4-positive T cells with anti-CD3 antibody increased the level of apoptosis twofold, and Fas–Fc blocked this effect substantially (Fig. 3c); DcR3–Fc blocked the induction of apoptosis to a similar extent. Thus, DcR3 binding blocks apoptosis induction by FasL.

FasL-induced apoptosis is important in elimination of virus-infected cells and cancer cells by natural killer cells and cytotoxic T lymphocytes; an alternative mechanism involves perforin and granzymes[1,14-16]. Peripheral blood natural killer cells triggered marked cell death in Jurkat T leukaemia cells (Fig. 3d); DcR3–Fc and Fas–Fc each reduced killing of target cells from ~65% to ~30%, with half-maximal inhibition at $\sim 1\ \mu g\ ml^{-1}$; the residual killing was probably mediated by the perforin/granzyme pathway. Thus, DcR3 binding blocks FasL-dependent natural killer cell activity. Higher DcR3–Fc and Fas–Fc concentrations were required to block natural killer cell activity compared with those required to block soluble FasL activity, which is consistent with the greater potency of membrane-associated FasL compared with soluble FasL[17].

Given the role of immune cytotoxic cells in elimination of tumour cells and the fact that DcR3 can act as an inhibitor of FasL, we proposed that DcR3 expression might contribute to the ability of some tumours to escape immune-cytotoxic attack. As genomic amplification frequently contributes to tumorigenesis, we investigated whether the DcR3 gene is amplified in cancer. We analysed DcR3 gene-copy number by quantitative polymerase chain

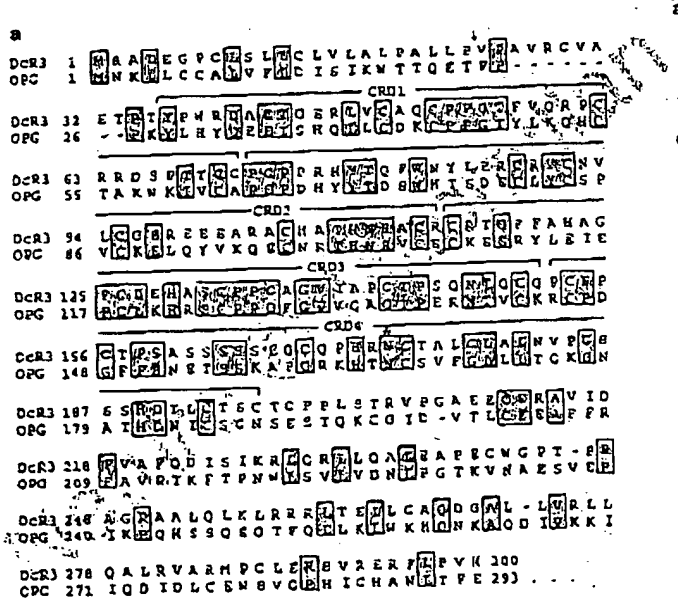


Figure 1 Primary structure and expression of human DcR3. a, Alignment of the amino-acid sequences of DcR3 and of osteoprotegerin (OPG); the C-terminal 101 residues of OPG are not shown. The putative signal cleavage site (arrow), the cysteine-rich domains (CRD 1–4), and the *N*-linked glycosylation site (asterisk) are shown. b, Expression of DcR3 mRNA. Northern hybridization analysis was done using the DcR3 cDNA as a probe and blots of poly(A)$^+$ RNA (Clontech) from human fetal and adult tissues or cancer cell lines. PBL, peripheral blood lymphocyte.

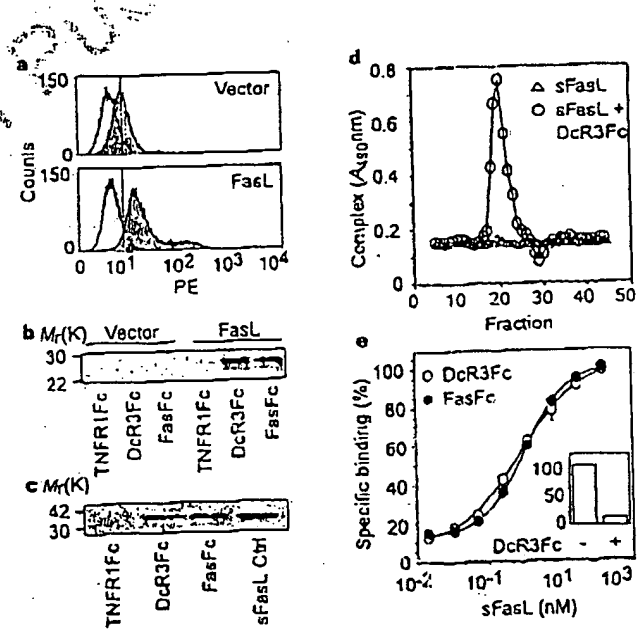


Figure 2 Interaction of DcR3 with FasL. a, 293 cells were transfected with pRK5 vector (top) or with pRK5 encoding full-length FasL (bottom), incubated with DcR3–Fc (solid line, shaded area), TNFR1–Fc (dotted line) or buffer control (dashed line) (the dashed and dotted lines overlap), and analysed for binding by FACS. Statistical analysis showed a significant difference ($P < 0.001$) between the binding of DcR3–Fc to cells transfected with FasL or pRK5. PE, phycoerythrin-labelled cells. b, 293 cells were transfected as in a and metabolically labelled, and cell supernatants were immunoprecipitated with Fc-tagged TNFR1, DcR3 or Fas. c, Purified soluble FasL (sFasL) was immunoprecipitated with TNFR1–Fc, DcR3–Fc or Fas–Fc and visualized by immunoblot with anti-FasL antibody. sFasL was loaded directly for comparison in the right-hand lane. d, Flag-tagged sFasL was incubated with DcR3–Fc or with buffer and resolved by gel filtration; column fractions were analysed in an assay that detects complexes containing DcR3–Fc and sFasL–Flag. e, Equilibrium binding of DcR3–Fc or Fas–Fc to sFasL–Flag. Inset, competition of DcR3–Fc with Fas–Fc for binding to sFasL–Flag.

reaction (PCR)[18] in genomic DNA from 35 primary lung and colon tumours, relative to pooled genomic DNA from peripheral blood leukocytes (PBLs) of 10 healthy donors. Eight of 18 lung tumours and 9 of 17 colon tumours showed DcR3 gene amplification, ranging from 2- to 18-fold (Fig. 4a, b). To confirm this result, we analysed the colon tumour DNAs with three more, independent sets of DcR3-based PCR primers and probes; we observed nearly the same amplification (data not shown).

We then analysed DcR3 mRNA expression in primary tumour tissue sections by *in situ* hybridization. We detected DcR3 expression in 6 out of 15 lung tumours, 2 out of 2 colon tumours, 2 out of 5 breast tumours, and 1 out of 1 gastric tumour (data not shown). A section through a squamous-cell carcinoma of the lung is shown in Fig. 4c. DcR3 mRNA was localized to infiltrating malignant epithelium, but was essentially absent from adjacent stroma, indicating tumour-specific expression. Although the individual tumour specimens that we analysed for mRNA expression and gene amplification were different, the *in situ* hybridization results are consistent with the finding that the DcR3 gene is amplified frequently in tumours. SW480 colon carcinoma cells, which showed abundant DcR3 mRNA expression (Fig. 1b), also had marked DcR3 gene amplification, as shown by quantitative PCR (fourfold) and by Southern blot hybridization (fivefold) (data not shown).

If DcR3 amplification in cancer is functionally relevant, then DcR3 should be amplified more than neighbouring genomic regions that are not important for tumour survival. To test this, we mapped the human DcR3 gene by radiation-hybrid analysis; DcR3 showed linkage to marker AFM218xe7 (T160), which maps to chromosome position 20q13. Next, we isolated from a bacterial artificial chromosome (BAC) library a human genomic clone that carries DcR3, and sequenced the ends of the clone's insert. We then determined, from the nine colon tumours that showed twofold or greater amplification of DcR3, the copy number of the DcR3-flanking sequences (reverse and forward) from the BAC, and of seven genomic markers that span chromosome 20 (Fig. 4d). The DcR3-linked reverse marker showed an average amplification of roughly threefold, slightly less than the approximately fourfold amplification of DcR3; the other markers showed little or no amplification. These data indicate that DcR3 may be at the 'epicentre' of a distal chromosome 20 region that is amplified in colon cancer, consistent with the possibility that DcR3 amplification promotes tumour survival.

Our results show that DcR3 binds specifically to FasL and inhibits FasL activity. We did not detect DcR3 binding to several other TNF-ligand-family members; however, this does not rule out the possibility that DcR3 interacts with other ligands, as do some other TNFR family members, including OPG[2,19].

FasL is important in regulating the immune response; however, little is known about how FasL function is controlled. One mechanism involves the molecule cFLIP, which modulates apoptosis signalling downstream of Fas[20]. A second mechanism involves proteolytic shedding of FasL from the cell surface[17]. DcR3 competes with Fas for

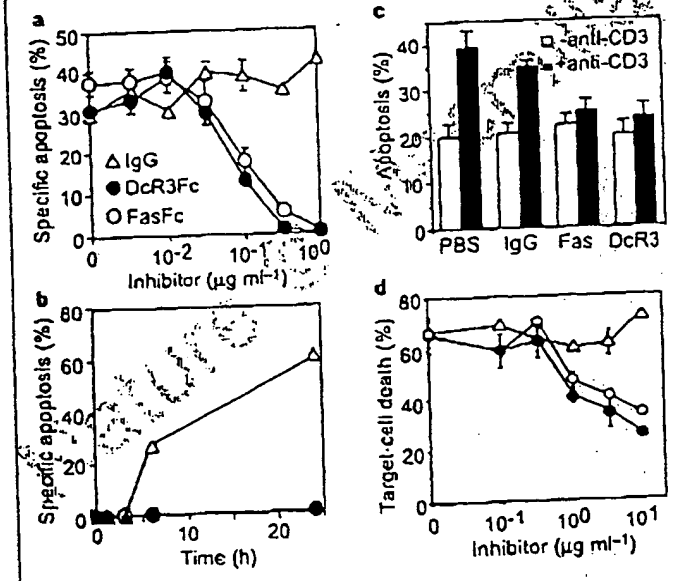


Figure 3 Inhibition of FasL activity by DcR3. a, Human Jurkat T leukaemia cells were incubated with Flag-tagged soluble FasL (sFasL; 5 ng ml^{-1}) oligomerized with anti-Flag antibody (0.1 μg ml^{-1}) in the presence of the proposed inhibitors DcR3-Fc, Fas-Fc or human IgG1 and assayed for apoptosis (mean ± s.e.m. of triplicates). b, Jurkat cells were incubated with sFasL–Flag plus anti-Flag antibody as in a, in presence of 1 μg ml^{-1} DcR3-Fc (filled circles), Fas-Fc (open circles) or human IgG1 (triangles), and apoptosis was determined at the indicated time points. c, Peripheral blood T cells were stimulated with PHA and interleukin-2, followed by control (white bars) or anti-CD3 antibody (filled bars), together with phosphate-buffered saline (PBS), human IgG1, Fas-Fc, or DcR3-Fc (10 μg ml^{-1}). After 16 h, apoptosis of CD4$^+$ cells was determined (mean ± s.e.m. of results from five donors). d, Peripheral blood natural killer cells were incubated with ^{51}Cr-labelled Jurkat cells in the presence of DcR3-Fc (filled circles), Fas-Fc (open circles) or human IgG1 (triangles), and target-cell death was determined by release of ^{51}Cr (mean ± s.d. for two donors, each in triplicate).

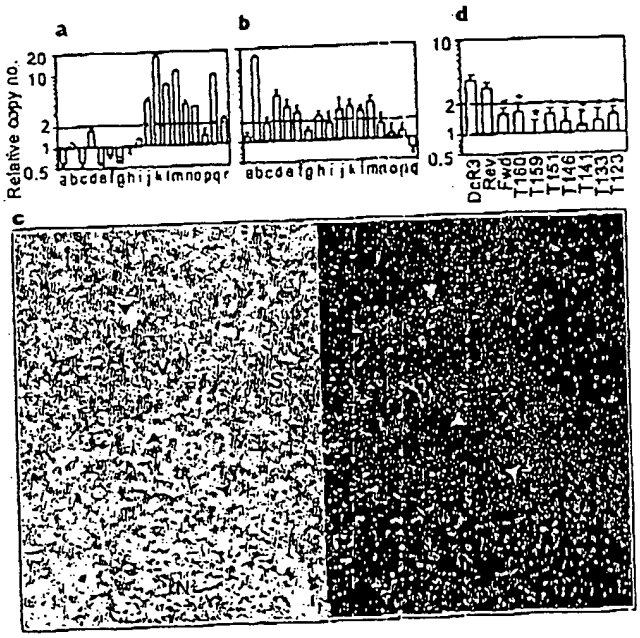


Figure 4 Genomic amplification of DcR3 in tumours. a, Lung cancers, comprising eight adenocarcinomas (c, d, f, g, h, j, k, r), seven squamous-cell carcinomas (a, e, m, n, o, p, q), one non-small-cell carcinoma (b), one small-cell carcinoma (i), and one bronchial adenocarcinoma (l). The data are means ± s.d. of 2 experiments done in duplicate. b, Colon tumours, comprising 17 adenocarcinomas. Data are means ± s.e.m. of five experiments done in duplicate. c, *In situ* hybridization analysis of DcR3 mRNA expression in a squamous-cell carcinoma of the lung. A representative bright-field image (left) and the corresponding dark-field image (right) show DcR3 mRNA over infiltrating malignant epithelium (arrowheads). Adjacent non-malignant stroma (S), blood vessel (V) and necrotic tumour tissue (N) are also shown. d, Average amplification of DcR3 compared with amplification of neighbouring genomic regions (reverse and forward, Rev and Fwd), the DcR3-linked marker T160, and other chromosome-20 markers. In the nine colon tumours showing DcR3 amplification of twofold or more (b). Data are from two experiments done in duplicate. Asterisk indicates $P < 0.01$ for a Student's *t*-test comparing each marker with DcR3.

FasL binding; hence, it may represent a third mechanism of extracellular regulation of FasL activity. A decoy receptor that modulates the function of the cytokine interleukin-1 has been described[21]. In addition, two decoy receptors that belong to the TNFR family, DcR1 and DcR2, regulate the FasL-related apoptosis-inducing molecule Apo2L[22]. Unlike DcR1 and DcR2, which are membrane-associated proteins, DcR3 is directly secreted into the extracellular space. One other secreted TNFR-family member is OPG[3], which shares greater sequence homology with DcR3 (31%) than do DcR1 (17%) or DcR2 (19%); OPG functions as a third decoy for Apo2L[19]. Thus, DcR3 and OPG define a new subset of TNFR-family members that function as secreted decoys to modulate ligands that induce apoptosis. Pox viruses produce soluble TNFR homologues that neutralize specific TNF-family ligands, thereby modulating the antiviral immune response[2]. Our results indicate that a similar mechanism, namely, production of a soluble decoy receptor for FasL, may contribute to immune evasion by certain tumours. □

Methods

Isolation of DcR3 cDNA. Several overlapping ESTs in GenBank (accession numbers AA025672, AA025673 and W67560) and in Lifeseq™ (Incyte Pharmaceuticals: accession numbers 1339238, 1533571, 1533650, 1542861, 1789372 and 2207027) showed similarity to members of the TNFR family. We screened human cDNA libraries by PCR with primers based on the region of EST consensus; fetal lung was positive for a product of the expected size. By hybridization to a PCR-generated probe based on the ESTs, one positive clone (DNA30942) was identified. When searching for potential alternatively spliced forms of DcR3 that might encode a transmembrane protein, we isolated 50 more clones; the coding regions of these clones were identical in size to that of the initial clone (data not shown).

Fc-fusion proteins (immunoadhesins). The entire DcR3 sequence, or the ectodomain of Fas or TNFR1, was fused to the hinge and Fc region of human IgG1, expressed in insect SF9 cells or in human 293 cells, and purified as described[23].

Fluorescence-activated cell sorting (FACS) analysis. We transfected 293 cells using calcium phosphate or Effectene (Qiagen) with pRK5 vector or pRK5 encoding full-length human FasL (2 μg), together with pRK5 encoding CrmA (2 μg) to prevent cell death. After 16 h, the cells were incubated with biotinylated DcR3–Fc or TNFR1–Fc and then with phycoerythrin-conjugated streptavidin (GibcoBRL), and were assayed by FACS. The data were analysed by Kolmogorov–Smirnov statistical analysis. There was some detectable staining of vector-transfected cells by DcR3–Fc; as these cells express little FasL (data not shown), it is possible that DcR3 recognized some other factor that is expressed constitutively on 293 cells.

Immunoprecipitation. Human 293 cells were transfected as above, and metabolically labelled with [^{35}S]cysteine and [^{35}S]methionine (0.5 mCi; Amersham). After 16 h of culture in the presence of z-VAD-fmk (10 μM), the medium was immunoprecipitated with DcR3–Fc, Fas–Fc or TNFR1–Fc (5 μg), followed by protein A–Sepharose (Repligen). The precipitates were resolved by SDS–PAGE and visualized on a phosphorimager (Fuji BAS2000). Alternatively, purified, Flag-tagged soluble FasL (1 μg) (Alexis) was incubated with each Fc-fusion protein (1 μg), precipitated with protein A–Sepharose, resolved by SDS–PAGE and visualized by immunoblotting with rabbit anti-FasL antibody (Oncogene Research).

Analysis of complex formation. Flag-tagged soluble FasL (25 μg) was incubated with buffer or with DcR3–Fc (40 μg) for 1.5 h at 24 °C. The reaction was loaded onto a Superdex 200 HR 10/30 column (Pharmacia) and developed with PBS; 0.6-ml fractions were collected. The presence of DcR3–Fc–FasL complex in each fraction was analysed by placing 100 μl aliquots into microtitre wells precoated with anti-human IgG (Boehringer) to capture DcR3–Fc, followed by detection with biotinylated anti-Flag antibody Bio M2 (Kodak) and streptavidin–horseradish peroxidase (Amersham). Calibration of the column indicated an apparent relative molecular mass of the complex of 420K (data not shown), which is consistent with a stoichiometry of two DcR3–Fc homodimers to two soluble FasL homotrimers.

Equilibrium binding analysis. Microtitre wells were coated with anti-human IgG, blocked with 2% BSA in PBS. DcR3–Fc or Fas–Fc was added, followed by serially diluted Flag-tagged soluble FasL. Bound ligand was detected with anti-Flag antibody as above. In the competition assay, Fas–Fc was immobilized as above, and the wells were blocked with excess IgG1 before addition of Flag-tagged soluble FasL plus DcR3–Fc.

T-cell AICD. CD3$^+$ lymphocytes were isolated from peripheral blood of individual donors using anti-CD3 magnetic beads (Miltenyi Biotech), stimulated with phytohaemagglutinin (PHA; 2 μg ml^{-1}) for 24 h, and cultured in the presence of interleukin-2 (100 U ml^{-1}) for 5 days. The cells were plated in wells coated with anti-CD3 antibody (Pharmingen) and analysed for apoptosis 16 h later by FACS analysis of annexin-V-binding of CD4$^+$ cells[24].

Natural killer cell activity. Natural killer cells were isolated from peripheral blood of individual donors using anti-CD56 magnetic beads (Miltenyi Biotech), and incubated for 16 h with ^{51}Cr-loaded Jurkat cells at an effector-to-target ratio of 1:1 in the presence of DcR3–Fc, Fas–Fc or human IgG1. Target-cell death was determined by release of ^{51}Cr in effector–target co-cultures relative to release of ^{51}Cr by detergent lysis of equal numbers of Jurkat cells.

Gene-amplification analysis. Surgical specimens were provided by J. Kern (lung tumours) and P. Quirke (colon tumours). Genomic DNA was extracted (Qiagen) and the concentration was determined using Hoechst dye 33258 intercalation fluorometry. Amplification was determined by quantitative PCR[18] using a TaqMan instrument (ABI). The method was validated by comparison of PCR and Southern hybridization data for the Myc and HER-2 oncogenes (data not shown). Gene-specific primers and fluorogenic probes were designed on the basis of the sequence of DcR3 or of nearby regions identified on a BAC carrying the human DcR3 gene; alternatively, primers and probes were based on Stanford Human Genome Center marker AFM218xe7 (T160), which is linked to DcR3 (likelihood score = 5.4), SHGC-36268 (T159), the nearest available marker which maps to ~500 kilobases from T160, and five extra markers that span chromosome 20. The DcR3-specific primer sequences were 5′-CTTCTTCGCGCACGCTG-3′ and 5′-ATCACGCCGGCACCAG-3′ and the fluorogenic probe sequence was 5′-(FAM)-ACACGATGCGTGCTCCAAGCAGAAp-(TAMARA), where FAM is 5′-fluorescein phosphoramidite. Relative gene-copy numbers were derived using the formula $2^{(\Delta CT)}$, where ΔCT is the difference in amplification cycles required to detect DcR3 in peripheral blood lymphocyte DNA compared to test DNA.

Received 24 September; accepted 5 November 1998.

1. Nagata, S. Apoptosis by death factor. *Cell* 88, 355–365 (1997).
2. Smith, C. A., Farrah, T. & Goodwin, R. G. The TNF receptor superfamily of cellular and viral proteins: activation, costimulation, and death. *Cell* 76, 959–962 (1994).
3. Simonet, W. S. *et al.* Osteoprotegerin: a novel secreted protein involved in the regulation of bone density. *Cell* 89, 309–319 (1997).
4. Suda, T., Takahashi, T., Golstein, P. & Nagata, S. Molecular cloning and expression of Fas ligand, a novel member of the TNF family. *Cell* 75, 1169–1178 (1993).
5. Pennica, D. *et al.* Human tumour necrosis factor: precursor structure, expression and homology to lymphotoxin. *Nature* 312, 724–729 (1984).
6. Pitti, R. M. *et al.* Induction of apoptosis by Apo-2 ligand, a new member of the tumor necrosis factor receptor family. *J. Biol. Chem.* 271, 12687–12690 (1996).
7. Wiley, S. R. *et al.* Identification and characterization of a new member of the TNF family that induces apoptosis. *Immunity* 3, 673–682 (1995).
8. Marsters, S. A. *et al.* Identification of a ligand for the death-domain-containing receptor Apo3. *Curr. Biol.* 8, 525–528 (1998).
9. Chicheportiche, Y. *et al.* TWEAK, a new secreted ligand in the TNF family that weakly induces apoptosis. *J. Biol. Chem.* 272, 32401–32410 (1997).
10. Wong, B. R. *et al.* TRANCE is a novel ligand of the TNFR family that activates c-Jun-N-terminal kinase in T cells. *J. Biol. Chem.* 272, 25190–25194 (1997).
11. Anderson, D. M. *et al.* A homolog of the TNF receptor and its ligand enhance T-cell growth and dendritic-cell function. *Nature* 390, 175–179 (1997).
12. Lacey, D. L. *et al.* Osteoprotegerin ligand is a cytokine that regulates osteoclast differentiation and activation. *Cell* 93, 165–176 (1998).
13. Dhein, J., Walczak, H., Baumler, C., Debatin, K. M. & Krammer, P. H. Autocrine T-cell suicide mediated by Apo1/(Fas/CD95). *Nature* 373, 438–441 (1995).
14. Arase, H., Arase, N. & Saito, T. Fas-mediated cytotoxicity by freshly isolated natural killer cells. *J. Exp. Med.* 181, 1235–1238 (1995).
15. Medvedev, A. E. *et al.* Regulation of Fas and Fas ligand expression in NK cells by cytokines and the involvement of Fas ligand in NK/LAK cell-mediated cytotoxicity. *Cytokine* 9, 394–404 (1997).
16. Moretta, A. Mechanisms in cell-mediated cytotoxicity. *Cell* 90, 13–18 (1997).
17. Tanaka, M., Itai, T., Adachi, M. & Nagata, S. Downregulation of Fas ligand by shedding. *Nature Med* 4, 31–36 (1998).
18. Gelmini, S. *et al.* Quantitative PCR-based homogeneous assay with fluorogenic probes to measure c-erbB-2 oncogene amplification. *Clin. Chem.* 43, 752–756 (1997).
19. Emery, J. G. *et al.* Osteoprotegerin is a receptor for the cytotoxic ligand TRAIL. *J. Biol. Chem.* 273, 14363–14367 (1998).
20. Wallach, D. Placing death under control. *Nature* 388, 123–125 (1997).
21. Colotta, F. *et al.* Interleukin-1 type II receptor: a decoy target for IL-1 that is regulated by IL-4. *Science* 261, 472–475 (1993).

22. Ashkenazi, A. & Dixit, V. M. Death receptors: signaling and modulation. *Science* 281, 1305–1308 (1998).
23. Ashkenazi, A. & Chamow, S. M. Immunoadhesins as research tools and therapeutic agents. *Curr. Opin. Immunol.* 9, 195–200 (1997).
24. Marsters, S. et al. Activation of apoptosis by Apo-2 ligand is independent of FADD but blocked by CrmA. *Curr. Biol.* 6, 750–752 (1996).

Acknowledgements. We thank C. Clark, D. Pennica and V. Dixit for comments, and J. Kern and P. Quirke for tumour specimens.

Correspondence and requests for materials should be addressed to A.A. (e-mail: aa@gene.com). The GenBank accession number for the DcR3 cDNA sequence is AF104419.

Crystal structure of the ATP-binding subunit of an ABC transporter

Li-Wei Hung*, Iris Xiaoyan Wang†, Kishiko Nikaido†, Pei-Qi Liu†, Giovanna Ferro-Luzzi Ames† & Sung-Hou Kim*‡

*E. O. Lawrence Berkeley National Laboratory, † Department of Molecular and Cell Biology, and ‡ Department of Chemistry, University of California at Berkeley, Berkeley, California 94720, USA

ABC transporters (also known as traffic ATPases) form a large family of proteins responsible for the translocation of a variety of compounds across membranes of both prokaryotes and eukaryotes[1]. The recently completed *Escherichia coli* genome sequence revealed that the largest family of paralogous *E. coli* proteins is composed of ABC transporters[2]. Many eukaryotic proteins of medical significance belong to this family, such as the cystic fibrosis transmembrane conductance regulator (CFTR), the P-glycoprotein (or multidrug-resistance protein) and the heterodimeric transporter associated with antigen processing (Tap1–Tap2). Here we report the crystal structure at 1.5 Å resolution of HisP, the ATP-binding subunit of the histidine permease, which is an ABC transporter from *Salmonella typhimurium*. We correlate the details of this structure with the biochemical, genetic and biophysical properties of the wild-type and several mutant HisP proteins. The structure provides a basis for understanding properties of ABC transporters and of defective CFTR proteins.

ABC transporters contain four structural domains: two nucleotide-binding domains (NBDs), which are highly conserved throughout the family, and two transmembrane domains[1]. In prokaryotes these domains are often separate subunits which are assembled into a membrane-bound complex; in eukaryotes the domains are generally fused into a single polypeptide chain. The periplasmic histidine permease of *S. typhimurium* and *E. coli*[1,3–8] is a well-characterized ABC transporter that is a good model for this superfamily. It consists of a membrane-bound complex, $HisQMP_2$, which comprises integral membrane subunits, HisQ and HisM, and two copies of HisP, the ATP-binding subunit. HisP, which has properties intermediate between those of integral and peripheral membrane proteins[9], is accessible from both sides of the membrane, presumably by its interaction with HisQ and HisM[6]. The two HisP subunits form a dimer, as shown by their cooperativity in ATP hydrolysis[5], the requirement for both subunits to be present for activity[6], and the formation of a HisP dimer upon chemical cross-linking. Soluble HisP also forms a dimer[7]. HisP has been purified and characterized in an active soluble form[7] which can be reconstituted into a fully active membrane-bound complex[8].

The overall shape of the crystal structure of the HisP monomer is that of an 'L' with two thick arms (arm I and arm II); the ATP-binding pocket is near the end of arm I (Fig. 1). A six-stranded β-sheet (β3 and β8–β12) spans both arms of the L, with a domain of a α- plus β-type structure (β1, β2, β4–β7, α1 and α2) on one side (within arm I) and a domain of mostly α-helices (α3–α9) on the

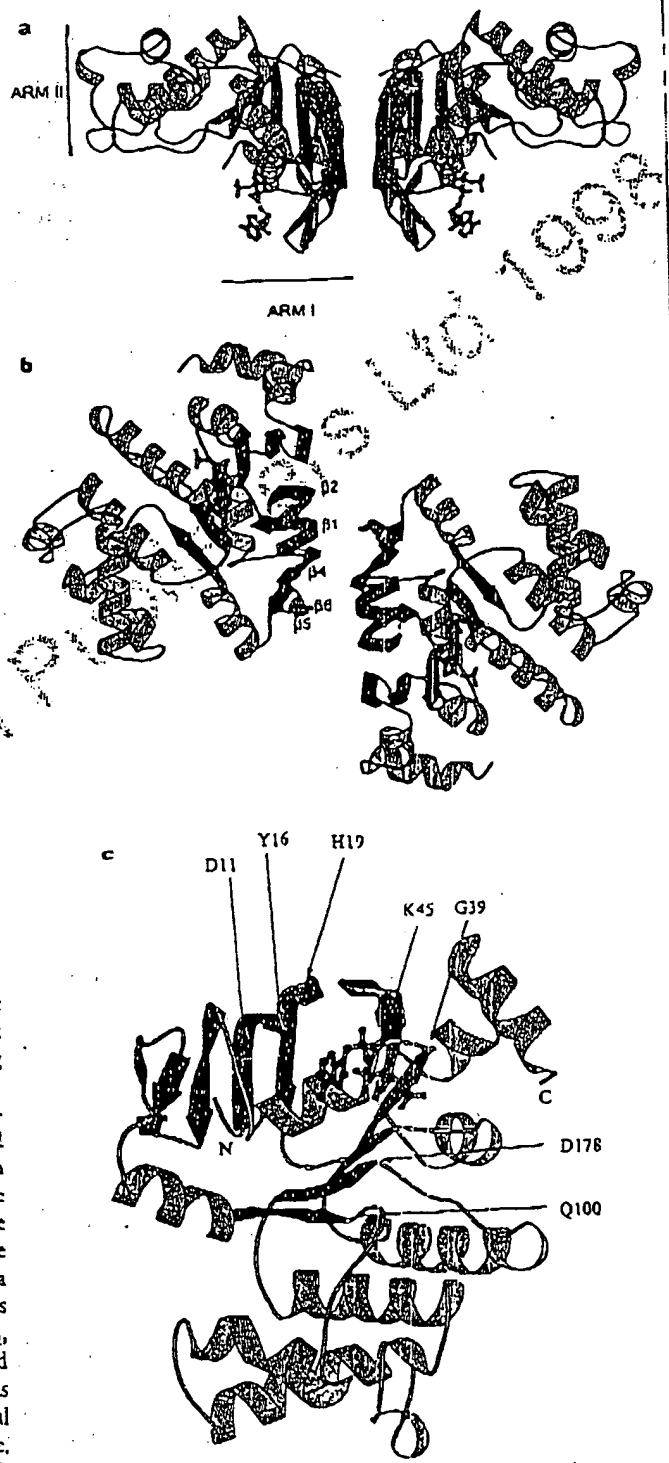


Figure 1 Crystal structure of HisP. a, View of the dimer along an axis perpendicular to its two-fold axis. The top and bottom of the dimer are suggested to face towards the periplasmic and cytoplasmic sides, respectively (see text). The thickness of arm II is about 25 Å, comparable to that of membrane. α-Helices are shown in orange and β-sheets in green. b, View along the two-fold axis of the HisP dimer, showing the relative displacement of the monomers not apparent in a. The β-strands at the dimer interface are labelled. c, View of one monomer from the bottom of arm I, as shown in a, towards arm II, showing the ATP-binding pocket. a–c, The protein and the bound ATP are in 'ribbon' and 'ball-and-stick' representations, respectively. Key residues discussed in the text are indicated in c. These figures were prepared with MOLSCRIPT[28]. N, amino terminus; C, C terminus.

HellerEhrman
ATTORNEYS

Int. J. Cancer: 78, 661–666 (1998)

Publication of the International Union Against Cancer
Publication de l'Union Internationale Contre le Cancer

NOVEL APPROACH TO QUANTITATIVE POLYMERASE CHAIN REACTION USING REAL-TIME DETECTION: APPLICATION TO THE DETECTION OF GENE AMPLIFICATION IN BREAST CANCER

Ivan BIÈCHE[1,2], Martine OLIVI[1], Marie-Hélène CHAMPÈME[2], Dominique VIDAUD[1], Rosette LIDEREAU[2] and Michel VIDAUD[1]*

[1]*Laboratoire de Génétique Moléculaire, Faculté des Sciences Pharmaceutiques et Biologiques de Paris, Paris, France*
[2]*Laboratoire d'Oncogénétique, Centre René Huguenin, St-Cloud, France*

Gene amplification is a common event in the progression of human cancers, and amplified oncogenes have been shown to have diagnostic, prognostic and therapeutic relevance. A kinetic quantitative polymerase-chain-reaction (PCR) method, based on fluorescent TaqMan methodology and a new instrument (ABI Prism 7700 Sequence Detection System) capable of measuring fluorescence in real-time, was used to quantify gene amplification in tumor DNA. Reactions are characterized by the point during cycling when PCR amplification is still in the exponential phase, rather than the amount of PCR product accumulated after a fixed number of cycles. None of the reaction components is limited during the exponential phase, meaning that values are highly reproducible in reactions starting with the same copy number. This greatly improves the precision of DNA quantification. Moreover, real-time PCR does not require post-PCR sample handling, thereby preventing potential PCR-product carry-over contamination; it possesses a wide dynamic range of quantification and results in much faster and higher sample throughput. The real-time PCR method, was used to develop and validate a simple and rapid assay for the detection and quantification of the 3 most frequently amplified genes (*myc*, *ccnd1* and *erbB2*) in breast tumors. Extra copies of *myc*, *ccnd1* and *erbB2* were observed in 10, 23 and 15%, respectively, of 108 breast-tumor DNA; the largest observed numbers of gene copies were 4.6, 18.6 and 15.1, respectively. These results correlated well with those of Southern blotting. The use of this new semi-automated technique will make molecular analysis of human cancers simpler and more reliable, and should find broad applications in clinical and research settings. *Int. J. Cancer* 78:661–666, 1998.
© 1998 Wiley-Liss, Inc.

Gene amplification plays an important role in the pathogenesis of various solid tumors, including breast cancer, probably because over-expression of the amplified target genes confers a selective advantage. The first technique used to detect genomic amplification was cytogenetic analysis. Amplification of several chromosome regions, visualized either as extrachromosomal double minutes (dmins) or as integrated homogeneously staining regions (HSRs), are among the main visible cytogenetic abnormalities in breast tumors. Other techniques such as comparative genomic hybridization (CGH) (Kallioniemi *et al.*, 1994) have also been used in broad searches for regions of increased DNA copy numbers in tumor cells, and have revealed some 20 amplified chromosome regions in breast tumors. Positional cloning efforts are underway to identify the critical gene(s) in each amplified region. To date, genes known to be amplified frequently in breast cancers include *myc* (8q24), *ccnd1* (11q13), and *erbB2* (17q12-q21) (for review, see Bièche and Lidereau, 1995).

Amplification of the *myc*, *ccnd1*, and *erbB2* proto-oncogenes should have clinical relevance in breast cancer, since independent studies have shown that these alterations can be used to identify sub-populations with a worse prognosis (Berns *et al.*, 1992; Schuuring *et al.*, 1992; Slamon *et al.*, 1987). Muss *et al.* (1994) suggested that these gene alterations may also be useful for the prediction and assessment of the efficacy of adjuvant chemotherapy and hormone therapy.

However, published results diverge both in terms of the frequency of these alterations and their clinical value. For instance, over 500 studies in 10 years have failed to resolve the controversy surrounding the link suggested by Slamon *et al.* (1987) between *erbB2* amplification and disease progression. These discrepancies are partly due to the clinical, histological and ethnic heterogeneity of breast cancer, but technical considerations are also probably involved.

Specific genes (DNA) were initially quantified in tumor cells by means of blotting procedures such as Southern and slot blotting. These batch techniques require large amounts of DNA (5–10 µg/reaction) to yield reliable quantitative results. Furthermore, meticulous care is required at all stages of the procedures to generate blots of sufficient quality for reliable dosage analysis. Recently, PCR has proven to be a powerful tool for quantitative DNA analysis, especially with minimal starting quantities of tumor samples (small, early-stage tumors and formalin-fixed, paraffin-embedded tissues).

Quantitative PCR can be performed by evaluating the amount of product either after a given number of cycles (end-point quantitative PCR) or after a varying number of cycles during the exponential phase (kinetic quantitative PCR). In the first case, an internal standard distinct from the target molecule is required to ascertain PCR efficiency. The method is relatively easy but implies generating, quantifying and storing an internal standard for each gene studied. Nevertheless, it is the most frequently applied method to date.

One of the major advantages of the kinetic method is its rapidity in quantifying a new gene, since no internal standard is required (an external standard curve is sufficient). Moreover, the kinetic method has a wide dynamic range (at least 5 orders of magnitude), giving an accurate value for samples differing in their copy number. Unfortunately, the method is cumbersome and has therefore been rarely used. It involves aliquot sampling of each assay mix at regular intervals and quantifying, for each aliquot, the amplification product. Interest in the kinetic method has been stimulated by a novel approach using fluorescent TaqMan methodology and a new instrument (ABI Prism 7700 Sequence Detection System) capable of measuring fluorescence in real time (Gibson *et al.*, 1996; Heid *et al.*, 1996). The TaqMan reaction is based on the 5′ nuclease assay first described by Holland *et al.* (1991). The latter uses the 5′ nuclease activity of Taq polymerase to cleave a specific fluorogenic oligonucleotide probe during the extension phase of PCR. The approach uses dual-labeled fluorogenic hybridization probes (Lee *et al.*, 1993). One fluorescent dye, co-valently linked to the 5′ end of the oligonucleotide, serves as a reporter [FAM (*i.e.*, 6-carboxy-fluorescein)] and its emission spectrum is quenched by a second fluorescent dye, TAMRA (*i.e.*, 6-carboxy-tetramethyl-rhodamine) attached to the 3′ end. During the extension phase of the PCR

Grant sponsors: Association Pour la Recherche sur le Cancer and Ministère de l'Enseignement Supérieur et de la Recherche.

*Correspondence to: Laboratoire de Génétique Moléculaire, Faculté des Sciences Pharmaceutiques et Biologiques de Paris, 4 Avenue de l'Observatoire, F-75006 Paris, France. Fax: (33)1-4407-1754.
E-mail: mvidaud@cesser.fr

Received 2 May 1998; Revised 30 June 1998

cycle, the fluorescent hybridization probe is hydrolyzed by the 5'-3' nucleolytic activity of DNA polymerase. Nuclease degradation of the probe releases the quenching of FAM fluorescence emission, resulting in an increase in peak fluorescence emission. The fluorescence signal is normalized by dividing the emission intensity of the reporter dye (FAM) by the emission intensity of a reference dye (*i.e.*, ROX, 6-carboxy-X-rhodamine) included in TaqMan buffer, to obtain a ratio defined as the Rn (normalized reporter) for a given reaction tube. The use of a sequence detector enables the fluorescence spectra of all 96 wells of the thermal cycler to be measured continuously during PCR amplification.

The real-time PCR method offers several advantages over other current quantitative PCR methods (Celi *et al.*, 1994): *(i)* the probe-based homogeneous assay provides a real-time method for detecting only specific amplification products, since specific hybridation of both the primers and the probe is necessary to generate a signal; *(ii)* the C_t (threshold cycle) value used for quantification is measured when PCR amplification is still in the log phase of PCR product accumulation. This is the main reason why C_t is a more reliable measure of the starting copy number than are end-point measurements, in which a slight difference in a limiting component can have a drastic effect on the amount of product; *(iii)* use of C_t values gives a wider dynamic range (at least 5 orders of magnitude), reducing the need for serial dilution; *(iv)* The real-time PCR method is run in a closed-tube system and requires no post-PCR sample handling, thus avoiding potential contamination; *(v)* the system is highly automated, since the instrument continuously measures fluorescence in all 96 wells of the thermal cycler during PCR amplification and the corresponding software processes, and analyzes the fluorescence data; *(vi)* the assay is rapid, as results are available just one minute after thermal cycling is complete; *(vii)* the sample throughput of the method is high, since 96 reactions can be analyzed in 2 hr.

Here, we applied this semi-automated procedure to determine the copy numbers of the 3 most frequently amplified genes in breast tumors (*myc, ccnd1* and *erb*B2), as well as 2 genes (*alb* and *app*) located in a chromosome region in which no genetic changes have been observed in breast tumors. The results for 108 breast tumors were compared with previous Southern-blot data for the same samples.

MATERIAL AND METHODS

Tumor and blood samples

Samples were obtained from 108 primary breast tumors removed surgically from patients at the Centre René Huguenin; none of the patients had undergone radiotherapy or chemotherapy. Immediately after surgery, the tumor samples were placed in liquid nitrogen until extraction of high-molecular-weight DNA. Patients were included in this study if the tumor sample used for DNA preparation contained more than 60% of tumor cells (histological analysis). A blood sample was also taken from 18 of the same patients.

DNA was extracted from tumor tissue and blood leukocytes according to standard methods.

Real-time PCR

Theoretical basis. Reactions are characterized by the point during cycling when amplification of the PCR product is first detected, rather than by the amount of PCR product accumulated after a fixed number of cycles. The higher the starting copy number of the genomic DNA target, the earlier a significant increase in fluorescence is observed. The parameter C_t (threshold cycle) is defined as the fractional cycle number at which the fluorescence generated by cleavage of the probe passes a fixed threshold above baseline. The target gene copy number in unknown samples is quantified by measuring C_t and by using a standard curve to determine the starting copy number. The precise amount of genomic DNA (based on optical density) and its quality (*i.e.*, lack of extensive degradation) are both difficult to assess. We therefore also quantified a control gene *(alb)* mapping to chromosome region 4q11-q13, in which no genetic alterations have been found in breast-tumor DNA by means of CGH (Kallioniemi *et al.*, 1994).

Thus, the ratio of the copy number of the target gene to the copy number of the *alb* gene normalizes the amount and quality of genomic DNA. The ratio defining the level of amplification is termed "N", and is determined as follows:

$$N = \frac{\text{copy number of target gene (}app, myc, ccnd1, erbB2\text{)}}{\text{copy number of reference gene (}alb\text{)}}.$$

Primers, probes, reference human genomic DNA and PCR consumables. Primers and probes were chosen with the assistance of the computer programs Oligo 4.0 (National Biosciences, Plymouth, MN), EuGene (Daniben Systems, Cincinnati, OH) and Primer Express (Perkin-Elmer Applied Biosystems, Foster City, CA).

Primers were purchased from DNAgency (Malvern, PA) and probes from Perkin-Elmer Applied Biosystems.

Nucleotide sequences for the oligonucleotide hybridization probes and primers are available on request.

The TaqMan PCR Core reagent kit, MicroAmp optical tubes, and MicroAmp caps were from Perkin-Elmer Applied Biosystems.

Standard-curve construction. The kinetic method requires a standard curve. The latter was constructed with serial dilutions of specific PCR products, according to Piatak *et al.* (1993). In practice, each specific PCR product was obtained by amplifying 20 ng of a standard human genomic DNA (Boehringer, Mannheim, Germany) with the same primer pairs as those used later for real-time quantitative PCR. The 5 PCR products were purified using MicroSpin S-400 HR columns (Pharmacia, Uppsala, Sweden) electrophorezed through an acrylamide gel and stained with ethidium bromide to check their quality. The PCR products were then quantified spectrophotometrically and pooled, and serially diluted 10-fold in mouse genomic DNA (Clontech, Palo Alto, CA) at a constant concentration of 2 ng/μl. The standard curve used for real-time quantitative PCR was based on serial dilutions of the pool of PCR products ranging from 10^{-7} (10^5 copies of each gene) to 10^{-10} (10^2 copies). This series of diluted PCR products was aliquoted and stored at −80°C until use.

The standard curve was validated by analyzing 2 known quantities of calibrator human genomic DNA (20 ng and 50 ng).

PCR amplification. Amplification mixes (50 μl) contained the sample DNA (around 20 ng, around 6600 copies of disomic genes), 10× TaqMan buffer (5 μl), 200 μM dATP, dCTP, dGTP, and 400 μM dUTP, 5 mM $MgCl_2$, 1.25 units of AmpliTaq Gold, 0.5 units of AmpErase uracil N-glycosylase (UNG), 200 nM each primer and 100 nM probe. The thermal cycling conditions comprised 2 min at 50°C and 10 min at 95°C. Thermal cycling consisted of 40 cycles at 95°C for 15 s and 65°C for 1 min. Each assay included: a standard curve (from 10^5 to 10^2 copies) in duplicate, a no-template control, 20 ng and 50 ng of calibrator human genomic DNA (Boehringer) in triplicate, and about 20 ng of unknown genomic DNA in triplicate (26 samples can thus be analyzed on a 96-well microplate). All samples with a coefficient of variation (CV) higher than 10% were retested.

All reactions were performed in the ABI Prism 7700 Sequence Detection System (Perkin-Elmer Applied Biosystems), which detects the signal from the fluorogenic probe during PCR.

Equipment for real-time detection. The 7700 system has a built-in thermal cycler and a laser directed via fiber optical cables to each of the 96 sample wells. A charge-coupled-device (CDD) camera collects the emission from each sample and the data are analyzed automatically. The software accompanying the 7700 system calculates C_t and determines the starting copy number in the samples.

Determination of gene amplification. Gene amplification was calculated as described above. Only samples with an N value higher than 2 were considered to be amplified.

RESULTS

To validate the method, real-time PCR was performed on genomic DNA extracted from 108 primary breast tumors, and 18 normal leukocyte DNA samples from some of the same patients. The target genes were the *myc, ccnd1* and *erb*B2 proto-oncogenes, and the β-amyloid precursor protein gene (*app*), which maps to a chromosome region (21q21.2) in which no genetic alterations have been found in breast tumors (Kallioniemi *et al.*, 1994). The reference disomic gene was the albumin gene (*alb*, chromosome 4q11-q13).

Validation of the standard curve and dynamic range of real-time PCR

The standard curve was constructed from PCR products serially diluted in genomic mouse DNA at a constant concentration of 2 ng/μl. It should be noted that the 5 primer pairs chosen to analyze the 5 target genes do not amplify genomic mouse DNA (data not shown). Figure 1 shows the real-time PCR standard curve for the *alb* gene. The dynamic range was wide (at least 4 orders of magnitude), with samples containing as few as 10^2 copies or as many as 10^5 copies.

*Copy-number ratio of the 2 reference genes (*app *and* alb*)*

The *app* to *alb* copy-number ratio was determined in 18 normal leukocyte DNA samples and all 108 primary breast-tumor DNA

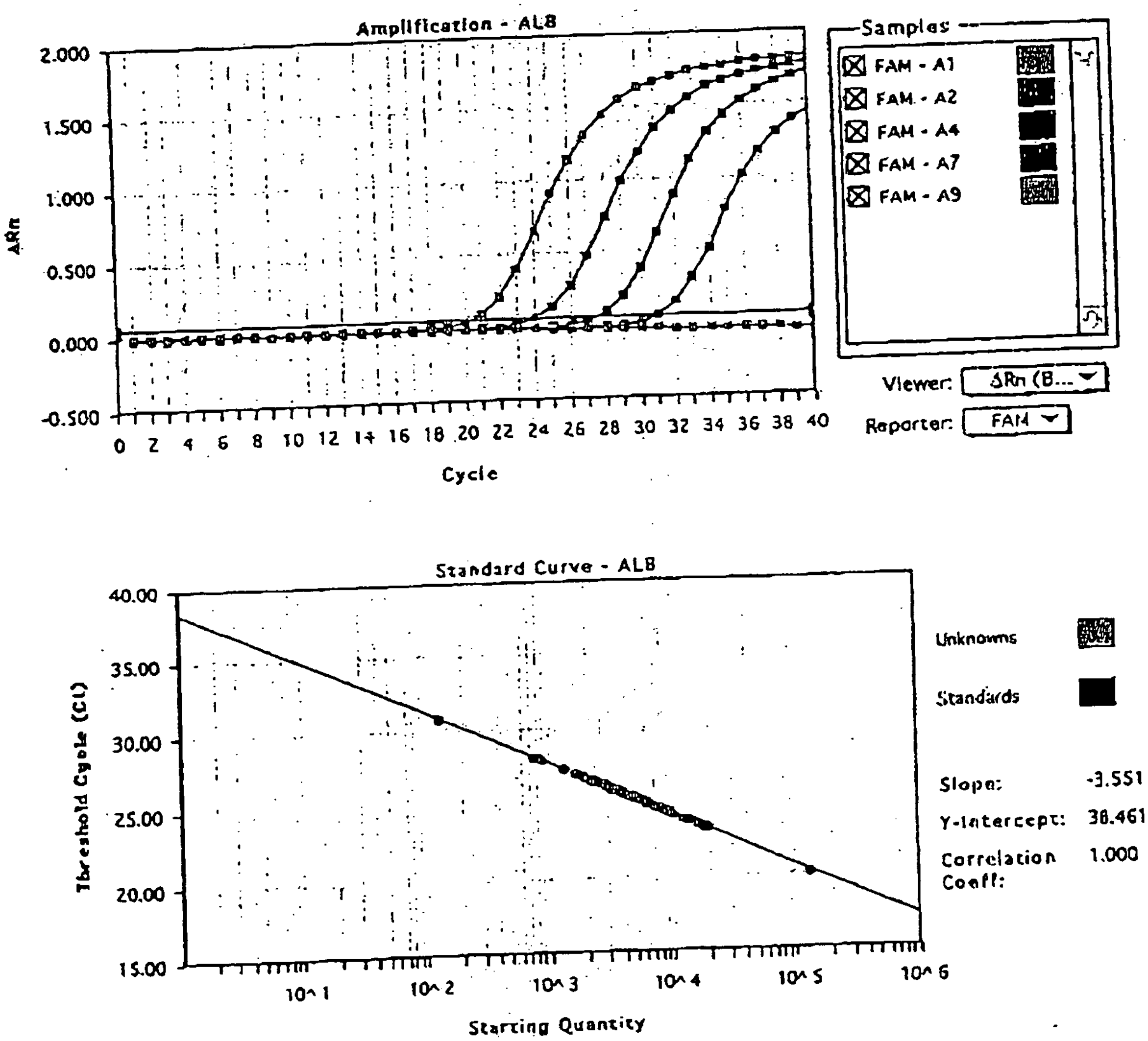


FIGURE 1 – Albumin (*alb*) gene dosage by real-time PCR. Top: Amplification plots for reactions with starting *alb* gene copy number ranging from 10^5 (A9), 10^4 (A7), 10^3 (A4) to 10^2 (A2) and a no-template control (A1). Cycle number is plotted *vs.* change in normalized reporter signal (ΔRn). For each reaction tube, the fluorescence signal of the reporter dye (FAM) is divided by the fluorescence signal of the passive reference dye (ROX), to obtain a ratio defined as the normalized reporter signal (Rn). ΔRn represents the normalized reporter signal (Rn) minus the baseline signal established in the first 15 PCR cycles. ΔRn increases during PCR as *alb* PCR product copy number increases until the reaction reaches a plateau. C_t (threshold cycle) represents the fractional cycle number at which a significant increase in Rn above a baseline signal (horizontal black line) can first be detected. Two replicate plots were performed for each standard sample, but the data for only one are shown here. Bottom: Standard curve plotting log starting copy number *vs.* C_t (threshold cycle). The black dots represent the data for standard samples plotted in duplicate and the red dots the data for unknown genomic DNA samples plotted in triplicate. The standard curve shows 4 orders of linear dynamic range.

samples. We selected these 2 genes because they are located in 2 chromosome regions (*app*, 21q21.2; *alb*, 4q11-q13) in which no obvious genetic changes (including gains or losses) have been observed in breast cancers (Kallioniemi *et al.*, 1994). The ratio for the 18 normal leukocyte DNA samples fell between 0.7 and 1.3 (mean 1.02 ± 0.21), and was similar for the 108 primary breast-tumor DNA samples (0.6 to 1.6, mean 1.06 ± 0.25), confirming that *alb* and *app* are appropriate reference disomic genes for breast-tumor DNA. The low range of the ratios also confirmed that the nucleotide sequences chosen for the primers and probes were not polymorphic, as mismatches of their primers or probes with the subject's DNA would have resulted in differential amplification.

myc, ccnd1 *and* erbB2 *gene dose in normal leukocyte DNA*

To determine the cut-off point for gene amplification in breast-cancer tissue, 18 normal leukocyte DNA samples were tested for the gene dose (N), calculated as described in "Material and Methods". The N value of these samples ranged from 0.5 to 1.3 (mean 0.84 ± 0.22) for *myc*, 0.7 to 1.6 (mean 1.06 ± 0.23) for *ccnd1* and 0.6 to 1.3 (mean 0.91 ± 0.19) for *erb*B2. Since N values for *myc*, *ccnd1* and *erb*B2 in normal leukocyte DNA consistently fell between 0.5 and 1.6, values of 2 or more were considered to represent gene amplification in tumor DNA.

myc, ccnd1 *and* erbB2 *gene dose in breast-tumor DNA*

myc, *ccnd1* and *erb*B2 gene copy numbers in the 108 primary breast tumors are reported in Table I. Extra copies of *ccnd1* were more frequent (23%, 25/108) than extra copies of *erb*B2 (15%, 16/108) and *myc* (10%, 11/108), and ranged from 2 to 18.6 for *ccnd1*, 2 to 15.1 for *erb*B2, and only 2 to 4.6 for the *myc* gene. Figure 2 and Table II represent tumors in which the *ccnd1* gene was amplified 16-fold (T145), 6-fold (T133) and non-amplified (T118). The 3 genes were never found to be co-amplified in the same tumor. *erb*B2 and *ccnd1* were co-amplified in only 3 cases, *myc* and *ccnd1* in 2 cases and *myc* and *erb*B2 in 1 case. This favors the hypothesis that gene amplifications are independent events in breast cancer. Interestingly, 5 tumors showed a decrease of at least 50% in the *erb*B2 copy number ($N < 0.5$), suggesting that they bore deletions of the 17q21 region (the site of *erb*B2). No such decrease in copy number was observed with the other 2 proto-oncogenes.

Comparison of gene dose determined by real-time quantitative PCR and Southern-blot analysis

Southern-blot analysis of *myc*, *ccnd1* and *erb*B2 amplifications had previously been done on the same 108 primary breast tumors. A perfect correlation between the results of real-time PCR and Southern blot was obtained for tumors with high copy numbers ($N \geq 5$). However, there were cases (1 *myc*, 6 *ccnd1* and 4 *erb*B2) in which real-time PCR showed gene amplification whereas Southern-blot did not, but these were mainly cases with low extra copy numbers (N from 2 to 2.9).

DISCUSSION

The clinical applications of gene amplification assays are currently limited, but would certainly increase if a simple, standardized and rapid method were perfected. Gene amplification status has been studied mainly by means of Southern blotting, but this method is not sensitive enough to detect low-level gene amplification nor accurate enough to quantify the full range of amplification values. Southern blotting is also time-consuming, uses radioactive reagents and requires relatively large amounts of high-quality genomic DNA, which means it cannot be used routinely in many laboratories. An amplification step is therefore required to determine the copy number of a given target gene from minimal quantities of tumor DNA (small early-stage tumors, cytopuncture specimens or formalin-fixed, paraffin-embedded tissues).

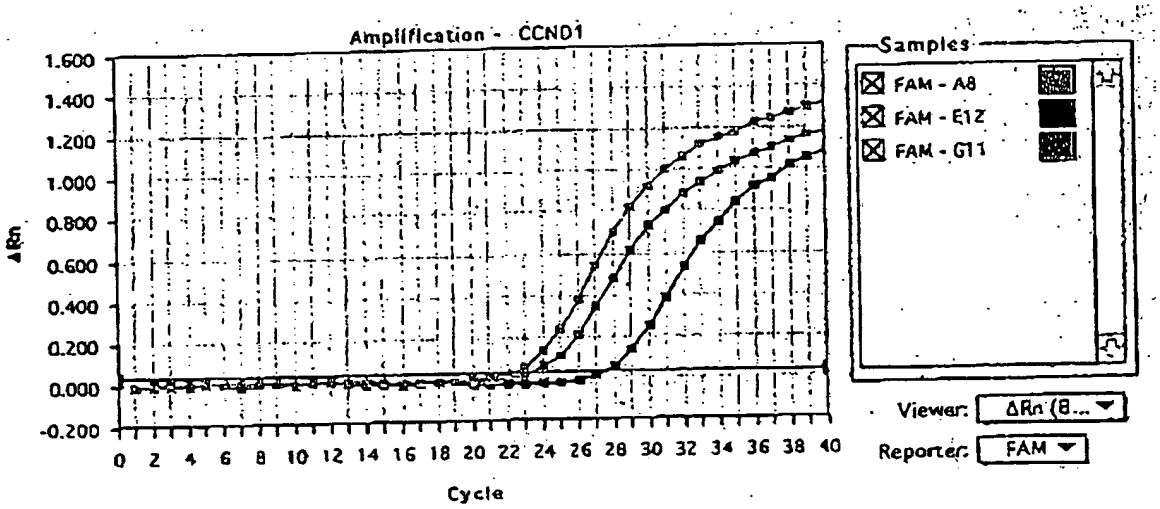
In this study, we validated a PCR method developed for the quantification of gene over-representation in tumors. The method, based on real-time analysis of PCR amplification, has several advantages over other PCR-based quantitative assays such as competitive quantitative PCR (Celi *et al.*, 1994). First, the real-time PCR method is performed in a closed-tube system, avoiding the risk of contamination by amplified products. Re-amplification of carryover PCR products in subsequent experiments can also be prevented by using the enzyme uracil N-glycosylase (UNG) (Longo *et al.*, 1990). The second advantage is the simplicity and rapidity of sample analysis, since no post-PCR manipulations are required. Our results show that the automated method is reliable. We found it possible to determine, in triplicate, the number of copies of a target gene in more than 100 tumors per day. Third, the system has a linear dynamic range of at least 4 orders of magnitude, meaning that samples do not have to contain equal starting amounts of DNA. This technique should therefore be suitable for analyzing formalin-fixed, paraffin-embedded tissues. Fourth, and above all, real-time PCR makes DNA quantification much more precise and reproducible, since it is based on C_t values rather than end-point measurement of the amount of accumulated PCR product. Indeed, the ABI Prism 7700 Sequence Detection System enables C_t to be calculated when PCR amplification is still in the exponential phase and when none of the reaction components is rate-limiting. The within-run CV of the C_t value for calibrator human DNA (5 replicates) was always below 5%, and the between-assay precision in 5 different runs was always below 10% (data not shown). In addition, the use of a standard curve is not absolutely necessary, since the copy number can be determined simply by comparing the C_t ratio of the target gene with that of reference genes. The results obtained by the 2 methods (with and without a standard curve) are similar in our experiments (data not shown). Moreover, unlike competitive quantitative PCR, real-time PCR does not require an internal control (the design and storage of internal controls and the validation of their amplification efficiency is laborious).

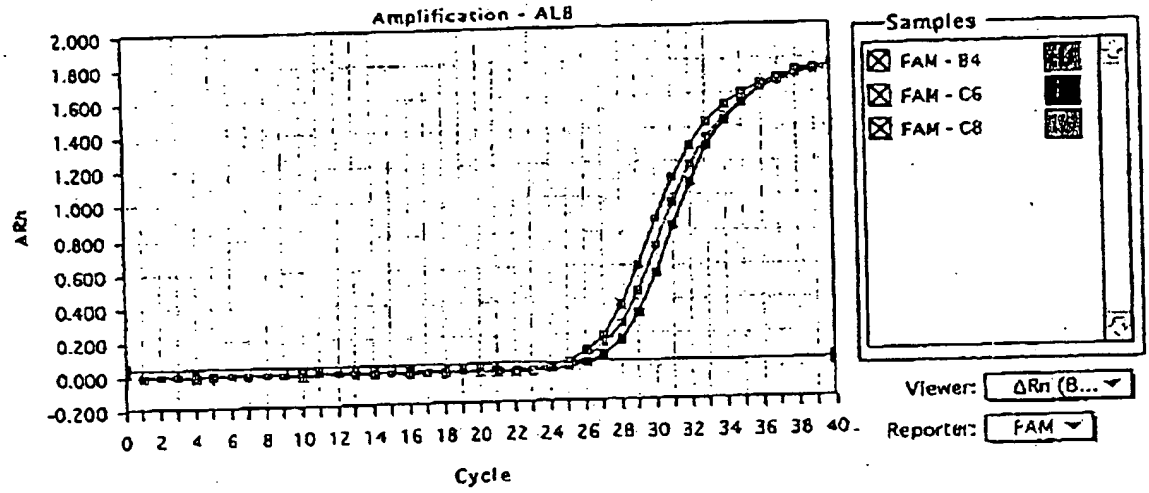
The only potential disavantage of real-time PCR, like all other PCR-based methods and solid-matrix blotting techniques (Southern blots and dot blots) is that is cannot avoid dilution artifacts inherent in the extraction of DNA from tumor cells contained in heterogeneous tissue specimens. Only FISH and immunohistochemistry can measure alterations on a cell-by-cell basis (Pauletti *et al.*, 1996; Slamon *et al.*, 1989). However, FISH requires expensive equipment and trained personnel and is also time-consuming. Moreover, FISH does not assess gene expression and therefore cannot detect cases in which the gene product is over-expressed in the absence of gene amplification, which will be possible in the future by real-time quantitative RT-PCR. Immunohistochemistry is subject to considerable variations in the hands of different teams, owing to alterations of target proteins during the procedure, the different primary antibodies and fixation methods used and the criteria used to define positive staining.

The results of this study are in agreement with those reported in the literature. (*i*) Chromosome regions 4q11-q13 and 21q21.2 (which bear *alb* and *app*, respectively) showed no genetic alterations in the breast-cancer samples studied here, in keeping with the results of CGH (Kallioniemi *et al.*, 1994). (*ii*) We found that amplifications of these 3 oncogenes were independent events, as reported by other teams (Berns *et al.*, 1992; Borg *et al.*, 1992). (*iii*) The frequency and degree of *myc* amplification in our breast tumor DNA series were lower than those of *ccnd1* and *erb*B2 amplification, confirming the findings of Borg *et al.* (1992) and Courjal *et al.* (1997). (*iv*) The maxima of *ccnd1* and *erb*B2 over-representation were 18-fold and 15-fold, also in keeping with earlier results (about

TABLE I – DISTRIBUTION OF AMPLIFICATION LEVEL (N) FOR *myc*, *ccnd1* AND *erb*B2 GENES IN 108 HUMAN BREAST TUMORS

Gene	Amplification level (N)			
	<0.5	0.5–1.9	2–4.9	≥5
myc	0	97 (89.8%)	11 (10.2%)	0
ccnd1	0	83 (76.9%)	17 (15.7%)	8 (7.4%)
*erb*B2	5 (4.6%)	87 (80.6%)	8 (7.4%)	8 (7.4%)





Tumor	CCND1 C_t	CCND1 Copy number	ALB C_t	ALB Copy number
T118	27.3	4605	26.5	4365
T133	23.2	61659	25.2	10092
T145	22.1	125892	25.6	7762

FIGURE 2 – *ccnd1* and alb gene dosage by real-time PCR in 3 breast tumor samples: T118 (E12, C6, black squares), T133 (G11, B4, red squares) and T145 (A8, C8, blue squares). Given the C_t of each sample, the initial copy number is inferred from the standard curve obtained during the same experiment. Triplicate plots were performed for each tumor sample, but the data for only one are shown here. The results are shown in Table II.

30-fold maximum) (Berns *et al.*, 1992; Borg *et al.*, 1992; Courjal *et al.*, 1997). (v) The *erb*B2 copy numbers obtained with real-time PCR were in good agreement with data obtained with other quantitative PCR-based assays in terms of the frequency and degree of amplification (An *et al.*, 1995; Deng *et al.*, 1996; Valeron *et al.*, 1996). Our results also correlate well with those recently published by Gelmini *et al.* (1997), who used the TaqMan system to measure *erb*B2 amplification in a small series of breast tumors (n = 25), but with an instrument (LS-50B luminescence spectrometer, Perkin-Elmer Applied Biosystems) which only allows end-

TABLE II – EXAMPLES OF *ccnd1* GENE DOSAGE RESULTS FROM 3 BREAST TUMORS[1]

Tumor	*ccnd1*			*alb*			Nccnd1/alb
	Copy number	Mean	SD	Copy number	Mean	SD	
T118	4525			4223			
	4605	4603	77	4365	4325	89	1.06
	4678			4387			
T133	59821			9787			
	61659	61100	1111	10092	10137	375	6.03
	61821			10533			
T145	128563			7321			
	125892	125392	3448	7762	7672	316	16.34
	121722			7933			

[1]For each sample, 3 replicate experiments were performed and the mean and the standard deviation (SD) was determined. The level of *ccnd1* gene amplification (*Nccnd1/alb*) is determined by dividing the average *ccnd1* copy number value by the average *alb* copy number value.

point measurement of fluorescence intensity. Here we report *myc* and *ccnd1* gene dosage in breast cancer by means of quantitative PCR. (vi) We found a high degree of concordance between real-time quantitative PCR and Southern blot analysis in terms of gene amplification, especially for samples with high copy numbers (≥5-fold). The slightly higher frequency of gene amplification (especially *ccnd1* and *erb*B2) observed by means of real-time quantitative PCR as compared with Southern-blot analysis may be explained by the higher sensitivity of the former method. However, we cannot rule out the possibility that some tumors with a few extra gene copies observed in real-time PCR had additional copies of an arm or a whole chromosome (trisomy, tetrasomy or polysomy) rather than true gene amplification. These 2 types of genetic alteration (polysomy and gene amplification) could be easily distinguished in the future by using an additional probe located on the same chromosome arm, but some distance from the target gene. It is noteworthy that high gene copy numbers have the greatest prognostic significance in breast carcinoma (Borg *et al.*, 1992; Slamon *et al.*, 1987).

Finally, this technique can be applied to the detection of gene deletion as well as gene amplification. Indeed, we found a decreased copy number of *erb*B2 (but not of the other 2 proto-oncogenes) in several tumors; *erb*B2 is located in a chromosome region (17q21) reported to contain both deletions and amplifications in breast cancer (Bièche and Lidereau, 1995).

In conclusion, gene amplification in various cancers can be used as a marker of pre-neoplasia, also for early diagnosis of cancer, staging, prognostication and choice of treatment. Southern blotting is not sufficiently sensitive, and FISH is lengthy and complex. Real-time quantitative PCR overcomes both these limitations, and is a sensitive and accurate method of analyzing large numbers of samples in a short time. It should find a place in routine clinical gene dosage.

ACKNOWLEDGEMENTS

RL is a research director at the Institut National de la Santé et de la Recherche Médicale (INSERM). We thank the staff of the Centre René Huguenin for assistance in specimen collection and patient care.

REFERENCES

AN, H.X., NIEDERACHER, D., BECKMANN, M.W., GÖHRING, U.J., SCHARL, A., PICARD, F., VAN ROEYEN, C., SCHNÜRCH, H.G. and BENDER, H.G., *erb*B2 gene amplification detected by fluorescent differential polymerase chain reaction in paraffin-embedded breast carcinoma tissues. *Int. J. Cancer (Pred. Oncol.)*, 64, 291–297 (1995).

BERNS, E.M.J.J., KLIJN, J.G.M., VAN PUTTEN, W.L.J., VAN STAVEREN, I.L., PORTENGEN, H. and FOEKENS, J.A., c-*myc* amplification is a better prognostic factor than HER2/*neu* amplification in primary breast cancer. *Cancer Res.*, 52, 1107–1113 (1992).

BIÈCHE, I. and LIDEREAU, R., Genetic alterations in breast cancer. *Genes Chrom. Cancer*, 14, 227–251 (1995).

BORG, A., BALDETORP, B., FERNO, M., OLSSON, H. and SIGURDSSON, H., c-*myc* amplification is an independent prognostic factor in post-menopausal breast cancer. *Int. J. Cancer*, 51, 687–691 (1992).

CELI, F.S., COHEN, M.M., ANTONARAKIS, S.E., WERTHEIMER, E., ROTH, J. and SHULOINER, A.R., Determination of gene dosage by a quantitative adaptation of the polymerase chain reaction (gd-PCR): rapid detection of deletions and duplications of gene sequences. *Genomics*, 21, 304–310 (1994).

COURJAL, F., CUNY, M., SIMONY-LAFONTAINE, J., LOUASSON, G., SPEISER, P., ZEILLINGER, R., RODRIGUEZ, C. and THEILLET, C., Mapping of DNA amplifications at 15 chromosomal localizations in 1875 breast tumors: definition of phenotypic groups. *Cancer Res.*, 57, 4360–4367 (1997).

DENG, G., YU, M., CHEN, L.C., MOORE, D., KURISU, W., KALLIONIEMI, A., WALDMAN, F.M., COLLINS, C. and SMITH, H.S., Amplifications of oncogene *erb*B-2 and chromosome 20q in breast cancer determined by differentially competitive polymerase chain reaction. *Breast Cancer Res. Treat.*, 40, 271–281 (1996).

GELMINI, S., ORLANDO, C., SESTINI, R., VONA, G., PINZANI, P., RUOCCO, L. and PAZZAGLI, M., Quantitative polymerase chain reaction-based homogeneous assay with fluorogenic probes to measure c-*erb*B-2 oncogene amplification. *Clin. Chem.*, 43, 752–758 (1997).

GIBSON, U.E.M., HEID, C.A. and WILLIAMS, P.M., A novel method for real-time quantitative RT-PCR. *Genome Res.*, 6, 995–1001 (1996).

HEID, C.A., STEVENS, J., LIVAK, K.J. and WILLIAMS, P.M., Real-time quantitative PCR. *Genome Res.*, 6, 986–994 (1996).

HOLLAND, P.M., ABRAMSON, R.D., WATSON, R. and GELFAND, D.H., Detection of specific polymerase chain reaction product by utilizing the 5' to 3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc. nat. Acad. Sci. (Wash.)*, 88, 7276–7280 (1991).

KALLIONIEMI, A., KALLIONIEMI, O.P., PIPER, J., TANNER, M., STOKKES, T., CHEN, L., SMITH, H.S., PINKEL, D., GRAY, J.W. and WALDMAN, F.M., Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization. *Proc. nat. Acad. Sci. (Wash.)*, 91, 2156–2160 (1994).

LEE, L.G., CONNELL, C.R. and BIOCH, W., Allelic discrimination by nick-translation PCR with fluorogenic probe. *Nucleic Acids Res.*, 21, 3761–3766 (1993).

LONGO, N., BERNINGER, N.S. and HARTLEY, J.L., Use of uracil DNA glycosylase to control carry-over contamination in polymerase chain reactions. *Gene*, 93, 125–128 (1990).

MUSS, H.B., THOR, A.D., BERRY, D.A., KUTE, T., LIU, E.T., KOERNER, P., CIRRINCIONE, C.T., BUDMAN, D.R., WOOD, W.C., BARCOS, M. and HENDERSON, I.C., c-*erb*B-2 expression and response to adjuvant therapy in women with node-positive early breast cancer. *New Engl. J. Med.*, 330, 1260–1266 (1994).

PAULETTI, G., GODOLPHIN, W., PRESS, M.F. and SALMON, D.J., Detection and quantification of HER-2/*neu* gene amplification in human breast cancer archival material using fluorescence *in situ* hybridization. *Oncogene*, 13, 63–72 (1996).

PIATAK, M., LUK, K.C., WILLIAMS, B. and LIFSON, J.D., Quantitative competitive polymerase chain reaction for accurate quantitation of HIV DNA and RNA species. *Biotechniques*, 14, 70–80 (1993).

SCHUURING, E., VERHOEVEN, E., VAN TINTEREN, H., PETERSE, J.L., NUNNIK, B., THUNNISSEN, F.B.J.M., DEVILEE, P., CORNELISSE, C.J., VAN DE VIJVER, M.J., MOOI, W.J. and MICHALIDES, R.J.A.M., Amplification of genes within the chromosome 11q13 region is indicative of poor prognosis in patients with operable breast cancer. *Cancer Res.*, 52, 5229–5234 (1992).

SLAMON, D.J., CLARK, G.M., WONG, S.G., LEVIN, W.S., ULLRICH, A. and MCGUIRE, W.L., Human breast cancer: correlation of relapse and survival with amplification of the HER-2/*neu* oncogene. *Science*, 235, 177–182 (1987).

SLAMON, D.J., GODOLPHIN, W., JONES, L.A., HOLT, J.A., WONG, S.G., KEITH, D.E., LEVIN, W.J., STUART, S.G., UDOVE, J., ULLRICH, A. and PRESS, M.F., Studies of the HER-2/*neu* proto-oncogene in human breast and ovarian cancer. *Science*, 244, 707–712 (1989).

VALERON, P.F., CHIRINO, R., FERNANDEZ, L., TORRES, S., NAVARRO, D., AGUIAR, J., CABRERA, J.J., DIAZ-CHICO, B.N. and DIAZ-CHICO, J.C., Validation of a differential PCR and an ELISA procedure in studying HER-2/*neu* status in breast cancer. *Int. J. Cancer*, 65, 129–133 (1996).

<first sequence: p1.DNA44804 (length = 598)
<second sequence: p1.holtzman (length = 673)

APPENDIX B

<597 matches in an overlap of 598: 99.83 percent similarity
<gaps in first sequence: 1 (75 residues), gaps in second sequence: 0
<score: 2895 (Dayhoff PAM 250 matrix, gap penalty = 8 + 4 per residue)
<endgaps not penalized

```
                    10        20        30        40        50        60
p1.DNA44804    MCSRVPLLLPLLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVF
               ************************************************************
p1.holtzman    MCSRVPLLLPLLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVF
                    10        20        30        40        50        60

                    70        80        90
p1.DNA44804    ENGITMLDASSFAGLPGLQLLDLSQNQIAS------------------------------
               *********.**********.*****
p1.holtzman    ENGITMLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETF
                    70        80        90       100       110       120

                                                                      100
p1.DNA44804    ---------------------------------------------LRLPRLLLLDLSHNS
                                                             ***************
p1.holtzman    RGLRRLERLYLGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNS
                         130       140       150       160       170       180

                 110       120       130       140       150       160
p1.DNA44804    LLALEPGILDTANVEALRLAGLGLQQLDEGLFSRLRNLHDLDVSDNQLERVPPVIRGLRG
               ************************************************************
p1.holtzman    LLALEPGILDTANVEALRLAGLGLQQLDEGLFSRLRNLHDLDVSDNQLERVPPVIRGLRG
                         190       200       210       220       230       240

                 170       180       190       200       210       220
p1.DNA44804    LTRLRLAGNTRIAQLRPEDLAGLAALQELDVSNLSLQALPGDLSGLFPRLRLLAAARNPF
               ************************************************************
p1.holtzman    LTRLRLAGNTRIAQLRPEDLAGLAALQELDVSNLSLQALPGDLSGLFPRLRLLAAARNPF
                         250       260       270       280       290       300

                 230       240       250       260       270       280
p1.DNA44804    NCVCPLSWFGPWVRESHVTLASPEETRCHFPPKNAGRLLLELDYADFGCPATTTTATVPT
               ************************************************************
p1.holtzman    NCVCPLSWFGPWVRESHVTLASPEETRCHFPPKNAGRLLLELDYADFGCPATTTTATVPT
                         310       320       330       340       350       360

                 290       300       310       320       330       340
p1.DNA44804    TRPVVREPTALSSSLAPTWLSPTAPATEAPSPPSTAPPTVGPVPQPQDCPPSTCLNGGTC
               ************************************************************
p1.holtzman    TRPVVREPTALSSSLAPTWLSPTAPATEAPSPPSTAPPTVGPVPQPQDCPPSTCLNGGTC
                         370       380       390       400       410       420

                 350       360       370       380       390       400
p1.DNA44804    HLGTRHHLACLCPEGFTGLYCESQMGQGTRPSPTPVTPRPPRSLTLGIEPVSPTSLRVGL
               ************************************************************
p1.holtzman    HLGTRHHLACLCPEGFTGLYCESQMGQGTRPSPTPVTPRPPRSLTLGIEPVSPTSLRVGL
                         430       440       450       460       470       480

                 410       420       430       440       450       460
p1.DNA44804    QRYLQGSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEYTVTQLRPNATYSVCVMPLGP
               ************************************************************
p1.holtzman    QRYLQGSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEYTVTQLRPNATYSVCVMPLGP
                         490       500       510       520       530       540

                 470       480       490       500       510       520
```

```
p1.DNA44804     GRVPEGE  CGEAHTPPAVHSNHAPVTQAREGNLP  IAPALAAVLLAALAAVGAAYCVR
                *******  **************************  **********************
p1.holtzman     GRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAVLLAALAAVGAAYCVR
                       550       560       570       580       590       600

                  530       540       550       560       570       580
p1.DNA44804     RGRAMAAAAQDKGQVGPGAGPLELEGVKVPLEPGPKATEGGGEALPSGSECEVPLMGFPG
                ************************************************************
p1.holtzman     RGRAMAAAAQDKGQVGPGAGPLELEGVKVPLEPGPKATEGGGEALPSGSECEVPLMGFPG
                       610       620       630       640       650       660

                  590
p1.DNA44804     PGLQSPLHAKPYI
                *************
p1.holtzman     PGLQSPLHAKPYI
                       670
```

Sequence file: /home/[illegible]y/va/Molbio/carpenda/temp.[illegible]tie/p1.holtzman
motifs in /usr/local/seq/libdata/motif.pro

Motif name: N-glycosylation site.
Accession: PS00001;
Motif: N[!P][ST][!P]

101 NLSN
117 NETF
273 NLSL
500 NLSG
528 NATY

Sequence file: /home/i y/va/Molbio/carpenda/temp. tie/p1.DNA44804
motifs in /usr/local/seq/libdata/motif.pro

Motif name: N-glycosylation site.
Accession: PS00001;
Motif: N[!P][ST][!P]

198 NLSL
425 NLSG
453 NATY

HMM file: /usr/seqdb/pfam/Pfam_ls
Sequence file: p1.DNA44804

- -

Query: DNA44804 [598 aa]

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
LRR	Leucine Rich Repeat	59.2	8.8e-14	7
LRRCT	Leucine rich repeat C-terminal domain	47.1	4e-10	1
EGF	EGF-like domain	30.0	5.4e-05	1
LRRNT	Leucine rich repeat N-terminal domain	29.8	6.5e-05	1
fn3	Fibronectin type III domain	13.0	0.15	1

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
LRRNT	1/1	23	51	..	1	31	[]	29.8	6.5e-05
LRR	1/7	53	76	..	1	25	[]	5.7	2.1e+02
LRR	2/7	77	102	..	1	25	[]	9.4	65
LRR	3/7	118	141	..	1	25	[]	10.4	44
LRR	4/7	142	164	..	1	25	[]	19.1	0.1
LRR	5/7	165	189	..	1	25	[]	11.1	26
LRR	6/7	190	212	..	1	25	[]	12.3	12
LRRCT	1/1	223	275	..	1	54	[]	47.1	4e-10
EGF	1/1	334	366	..	1	45	[]	30.0	5.4e-05
LRR	7/7	415	437	..	1	25	[]	3.1	4.8e+02
fn3	1/1	383	474	..	1	84	[]	13.0	0.15

HMM file: /usr/seqdb/pfam/Pfam_ls
Sequence file: p1.holtzman

- -

Query: holtzman [673 aa]

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
LRR	Leucine Rich Repeat	108.8	1e-28	11
LRRCT	Leucine rich repeat C-terminal domain	47.1	4e-10	1
EGF	EGF-like domain	30.0	5.4e-05	1
LRRNT	Leucine rich repeat N-terminal domain	29.8	6.5e-05	1
fn3	Fibronectin type III domain	13.0	0.15	1

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
LRRNT	1/1	23	51	..	1	31	[]	29.8	6.5e-05
LRR	1/11	53	76	..	1	25	[]	6.1	1.9e+02
LRR	2/11	77	100	..	1	25	[]	21.6	0.019
LRR	3/11	101	124	..	1	25	[]	15.6	1.2
LRR	4/11	125	148	..	1	25	[]	18.1	0.21
LRR	5/11	149	169	..	1	25	[]	9.7	58
LRR	6/11	170	192	..	1	25	[]	6.1	1.8e+02
LRR	7/11	193	216	..	1	25	[]	10.4	44
LRR	8/11	217	239	..	1	25	[]	19.1	0.1
LRR	9/11	240	264	..	1	25	[]	11.1	26
LRR	10/11	265	287	..	1	25	[]	12.3	12
LRRCT	1/1	298	350	..	1	54	[]	47.1	4e-10
EGF	1/1	409	441	..	1	45	[]	30.0	5.4e-05
LRR	11/11	490	512	..	1	25	[]	3.1	4.8e+02
fn3	1/1	458	549	..	1	84	[]	13.0	0.15

DECLARATION OF PAUL POLAKIS, Ph.D.

I, Paul Polakis, Ph.D., declare and say as follows:

1. I was awarded a Ph.D. by the Department of Biochemistry of the Michigan State University in 1984. My scientific Curriculum Vitae is attached to and forms part of this Declaration (Exhibit A).

2. I am currently employed by Genentech, Inc. where my job title is Staff Scientist. Since joining Genentech in 1999, one of my primary responsibilities has been leading Genentech's Tumor Antigen Project, which is a large research project with a primary focus on identifying tumor cell markers that find use as targets for both the diagnosis and treatment of cancer in humans.

3. As part of the Tumor Antigen Project, my laboratory has been analyzing differential expression of various genes in tumor cells relative to normal cells. The purpose of this research is to identify proteins that are abundantly expressed on certain tumor cells and that are either (i) not expressed, or (ii) expressed at lower levels, on corresponding normal cells. We call such differentially expressed proteins "tumor antigen proteins". When such a tumor antigen protein is identified, one can produce an antibody that recognizes and binds to that protein. Such an antibody finds use in the diagnosis of human cancer and may ultimately serve as an effective therapeutic in the treatment of human cancer.

4. In the course of the research conducted by Genentech's Tumor Antigen Project, we have employed a variety of scientific techniques for detecting and studying differential gene expression in human tumor cells relative to normal cells, at genomic DNA, mRNA and protein levels. An important example of one such technique is the well known and widely used technique of microarray analysis which has proven to be extremely useful for the identification of mRNA molecules that are differentially expressed in one tissue or cell type relative to another. In the course of our research using microarray analysis, we have identified approximately 200 gene transcripts that are present in human tumor cells at significantly higher levels than in corresponding normal human cells. To date, we have generated antibodies that bind to about 30 of the tumor antigen proteins expressed from these differentially expressed gene transcripts and have used these antibodies to quantitatively determine the level of production of these tumor antigen proteins in both human cancer cells and corresponding normal cells. We have then compared the levels of mRNA and protein in both the tumor and normal cells analyzed.

5. From the mRNA and protein expression analyses described in paragraph 4 above, we have observed that there is a strong correlation between changes in the level of mRNA present in any particular cell type and the level of protein

expressed from that mRNA in that cell type. In approximately 80% of our observations we have found that increases in the level of a particular mRNA correlates with changes in the level of protein expressed from that mRNA when human tumor cells are compared with their corresponding normal cells.

6. Based upon my own experience accumulated in more than 20 years of research, including the data discussed in paragraphs 4 and 5 above and my knowledge of the relevant scientific literature, it is my considered scientific opinion that for human genes, an increased level of mRNA in a tumor cell relative to a normal cell typically correlates to a similar increase in abundance of the encoded protein in the tumor cell relative to the normal cell. In fact, it remains a central dogma in molecular biology that increased mRNA levels are predictive of corresponding increased levels of the encoded protein. While there have been published reports of genes for which such a correlation does not exist, it is my opinion that such reports are exceptions to the commonly understood general rule that increased mRNA levels are predictive of corresponding increased levels of the encoded protein.

7. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information or belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful statements may jeopardize the validity of the application or any patent issued thereon.

Dated: 5/07/04

By: Paul Polakis

Paul Polakis, Ph.D.

SV 2031808 v1

CURRICULUM VITAE

PAUL G. POLAKIS
Staff Scientist
Genentech, Inc
1 DNA Way, MS#40
S. San Francisco, CA 94080

EDUCATION:

Ph.D., Biochemistry, Department of Biochemistry,
Michigan State University (1984)

B.S., Biology. College of Natural Science, Michigan State University (1977)

PROFESSIONAL EXPERIENCE:

2002-present	Staff Scientist, Genentech, Inc S. San Francisco, CA
1999- 2002	Senior Scientist, Genentech, Inc., S. San Francisco, CA
1997 -1999	Research Director Onyx Pharmaceuticals, Richmond, CA
1992- 1996	Senior Scientist, Project Leader, Onyx Pharmaceuticals, Richmond, CA
1991-1992	Senior Scientist, Chiron Corporation, Emeryville, CA.
1989-1991	Scientist, Cetus Corporation, Emeryville CA.
1987-1989	Postdoctoral Research Associate, Genentech, Inc., South SanFrancisco, CA.
1985-1987	Postdoctoral Research Associate, Department of Medicine, Duke University Medical Center, Durham, NC

1984-1985	Assistant Professor, Department of Chemistry, Oberlin College, Oberlin, Ohio
1980-1984	Graduate Research Assistant, Department of Biochemistry, Michigan State University East Lansing, Michigan

PUBLICATIONS:

1. **Polakis, P G.** and Wilson, J. E. 1982 Purification of a Highly Bindable Rat Brain Hexokinase by High Performance Liquid Chromatography. **Biochem. Biophys. Res. Commun**. 107, 937-943.

2. **Polakis, P.G.** and Wilson, J. E. 1984 Proteolytic Dissection of Rat Brain Hexokinase: Determination of the Cleavage Pattern during Limited Digestion with Trypsin. **Arch. Biochem. Biophys**. 234, 341-352.

3. **Polakis, P. G.** and Wilson, J. E. 1985 An Intact Hydrophobic N-Terminal Sequence is Required for the Binding Rat Brain Hexokinase to Mitochondria. **Arch. Biochem. Biophys**. 236, 328-337.

4. Uhing, R.J., **Polakis,P.G.** and Snyderman, R. 1987 Isolaton of GTP-binding Proteins from Myeloid HL60 Cells. **J. Biol. Chem**. 262, 15575-15579.

5. **Polakis, P.G.**, Uhing, R.J. and Snyderman, R. 1988 The Formylpeptide Chemoattractant Receptor Copurifies with a GTP-binding Protein Containing a Distinct 40 kDa Pertussis Toxin Substrate. **J. Biol. Chem**. 263, 4969-4979.

6. Uhing, R. J., Dillon, S., **Polakis, P. G.**, Truett, A. P. and Snyderman, R. 1988 Chemoattractant Receptors and Signal Transduction Processes in Cellular and Molecular Aspects of Inflammation (Poste, G. and Crooke, S. T. eds.) pp 335-379.

7. **Polakis, P.G.**, Evans, T. and Snyderman 1989 Multiple Chromatographic Forms of the Formylpeptide Chemoattractant Receptor and their Relationship to GTP-binding Proteins. **Biochem. Biophys. Res. Commun**. 161, 276-283.

8. **Polakis, P. G.**, Snyderman, R. and Evans, T. 1989 Characterization of G25K, a GTP-binding Protein Containing a Novel Putative Nucleotide Binding Domain. **Biochem. Biophys. Res. Comun.** 160, 25-32.

9. **Polakis, P.**, Weber,R.F., Nevins,B., Didsbury, J. Evans,T. and Snyderman, R. 1989 Identification of the ral and rac1 Gene Products, Low Molecular Mass GTP-binding Proteins from Human Platelets. **J. Biol. Chem**. 264, 16383-16389.

10. Snyderman, R., Perianin, A., Evans, T., **Polakis, P.** and Didsbury, J. 1989 G Proteins and Neutrophil Function. In ADP-Ribosylating Toxins and G Proteins: Insights into Signal Transduction. (J. Moss and M. Vaughn, eds.) Amer. Soc. Microbiol. pp. 295-323.

11. Hart, M.J., **Polakis, P.G.**, Evans, T. and Cerrione, R.A. 1990 The Identification and Charaterization of an Epidermal Growth Factor-Stimulated Phosphorylation of a Specific Low Molecular Mass GTP-binding Protein in a Reconstituted Phospholipid Vesicle System. **J. Biol. Chem**. 265, 5990-6001.

12. Yatani, A., Okabe, K., **Polakis, P.** Halenbeck, R. McCormick, F. and Brown, A. M. 1990 ras p21 and GAP Inhibit Coupling of Muscarinic Receptors to Atrial K^+ Channels. **Cell**. 61, 769-776.

13. Munemitsu, S., Innis, M.A., Clark, R., McCormick, F., Ullrich, A. and **Polakis, P.G.** 1990 Molecular Cloning and Expression of a G25K cDNA, the Human Homolog of the Yeast Cell Cycle Gene CDC42. **Mol. Cell. Biol**. 10, 5977-5982.

14. **Polakis, P.G.** Rubinfeld, B. Evans, T. and McCormick, F. 1991 Purification of Plasma Membrane-Associated GTPase Activating Protein (GAP) Specific for rap-1/krev-1 from HL60 Cells. **Proc. Natl. Acad. Sci. USA** 88, 239-243.

15. Moran, M. F., **Polakis, P.**, McCormick, F., Pawson, T. and Ellis, C. 1991 Protein Tyrosine Kinases Regulate the Phosphorylation, Protein Interactions, Subcellular Distribution, and Activity of p21ras GTPase Activating Protein. **Mol. Cell. Biol**. 11, 1804-1812

16. Rubinfeld, B., Wong, G., Bekesi, E. Wood, A. McCormick, F. and **Polakis, P. G.** 1991 A Synthetic Peptide Corresponding to a Sequence in the GTPase Activating Protein Inhibits $p21^{ras}$ Stimulation and Promotes Guanine Nucleotide Exchange. **Internatl. J. Peptide and Prot. Res**. 38, 47-53.

17. Rubinfeld, B., Munemitsu, S., Clark, R., Conroy, L., Watt, K., Crosier, W., McCormick, F., and **Polakis, P**. 1991 Molecular Cloning of a GTPase Activating Protein Specific for the Krev-1 Protein $p21^{rap1}$. **Cell** 65, 1033-1042.

18. Zhang, K. Papageorge, A., G., Martin, P., Vass, W. C., Olah, Z., **Polakis, P.**, McCormick, F. and Lowy, D, R. 1991 Heterogenous Amino Acids in RAS and Rap1A Specifying Sensitivity to GAP Proteins. **Science** 254, 1630-1634.

19. Martin, G., Yatani, A., Clark, R., **Polakis, P.**, Brown, A. M. and McCormick, F. 1992 GAP Domains Responsible for $p21^{ras}$-dependent Inhibition of Muscarinic Atrial K^+ Channel Currents. **Science** 255, 192-194.

20. McCormick, F., Martin, G. A., Clark, R., Bollag, G. and **Polakis, P** . 1992 Regulation of p21ras by GTPase Activating Proteins. Cold Spring Harbor **Symposia on Quantitative Biology**. Vol. 56, 237-241.

21. Pronk, G. B., **Polakis, P.**, Wong, G., deVries-Smits, A. M., Bos J. L. and McCormick, F. 1992 $p60^{v\text{-}src}$ Can Associate with and Phosphorylate the $p21^{ras}$ GTPase Activating Protein. **Oncogene** 7,389-394.

22. **Polakis P**. and McCormick, F. 1992 Interactions Between $p21^{ras}$ Proteins and Their GTPase Activating Proteins. In **Cancer Surveys** (Franks, L. M., ed.) 12, 25-42.

23. Wong, G., Muller, O., Clark, R., Conroy, L., Moran, M., **Polakis,P.** and McCormick, F. 1992 Molecular coloning and nucleic acid binding properties of the GAP-associated tyrosine phosphoprotein p62. **Cell** 69, 551-558.

24. Polakis, P., Rubinfeld, B. and McCormick, F. 1992 Phosphorylation of rap1GAP in vivo and by cAMP-dependent Kinase and the Cell Cycle $p34^{cdc2}$ Kinase in vitro. **J. Biol. Chem**. 267, 10780-10785.

25. McCabe, P.C., Haubrauck, H., **Polakis, P.**, McCormick, F., and Innis, M. A. 1992 Functional Interactions Between $p21^{rap1A}$ and Components of the Budding pathway of *Saccharomyces cerevisiae*. **Mol. Cell. Biol**. 12, 4084-4092.

26. Rubinfeld, B., Crosier, W.J., Albert, I., Conroy,L., Clark, R., McCormick, F. and **Polakis, P**. 1992 Localization of the rap1GAP Catalytic Domain and Sites of Phosphorylation by Mutational Analysis. **Mol. Cell . Biol**. 12, 4634-4642.

27. Ando, S., Kaibuchi, K., Sasaki, K., Hiraoka, T., Nishiyama, T., Mizuno, T., Asada, M., Nunoi, H., Matsuda, I., Matsuura, Y., **Polakis, P.**, McCormick, F. and Takai, Y. 1992 Post-translational processing of rac p21s is important both for their interaction with the GDP/GTP exchange proteins and for their activation of NADPH oxidase. **J. Biol. Chem**. 267, 25709-25713.

28. Janoueix-Lerosey, I., **Polakis, P.**, Tavitian, A. and deGunzberg, J. 1992 Regulation of the GTPase activity of the ras-related rap2 protein. **Biochem. Biophys. Res. Commun**. 189, 455-464.

29. Polakis, P. 1993 GAPs Specific for the rap1/Krev-1 Protein. in GTP-binding Proteins: the *ras*-superfamily. (J.C. LaCale and F. McCormick, eds.) 445-452.

30. Polakis, P. and McCormick, F. 1993 Structural requirements for the interaction of $p21^{ras}$ with GAP, exchange factors, and its bological effector target. **J. Biol Chem**. 268, 9157-9160.

31. Rubinfeld, B., Souza, B. Albert, I., Muller, O., Chamberlain, S., Masiarz, F., Munemitsu, S. and **Polakis, P.** 1993 Association of the APC gene product with beta- catenin. **Science** 262, 1731-1734.

32. Weiss, J., Rubinfeld, B., **Polakis, P.**, McCormick, F. Cavenee, W. A. and Arden, K. 1993 The gene for human rap1-GTPase activating protein (rap1GAP) maps to chromosome 1p35-1p36.1. **Cytogenet. Cell Genet**. 66, 18-21.

33. Sato, K. Y., **Polakis, P.**, Haubruck, H., Fasching, C. L., McCormick, F. and Stanbridge, E. J. 1994 Analysis of the tumor suppressor acitvity of the K-rev gene in human tumor cell lines. **Cancer Res**. 54, 552-559.

34. Janoueix-Lerosey, I., Fontenay, M., Tobelem, G., Tavitian, A., **Polakis, P**. and DeGunzburg, J. 1994 Phosphorylation of rap1GAP during the cell cycle. **Biochem. Biophys. Res. Commun**. 202, 967-975

35. Munemitsu, S., Souza, B., Mueller, O., Albert, I., Rubinfeld, B., and **Polakis, P.** 1994 The APC gene product associates with microtubules in vivo and affects their assembly in vitro. **Cancer Res**. 54, 3676-3681.

36. Rubinfeld, B. and **Polakis, P.** 1995 Purification of baculovirus produced rap1GAP. **Methods Enz.** 255,31

37. Polakis, P. 1995 Mutations in the APC gene and their implications for protein structure and function. **Current Opinions in Genetics and Development** 5, 66-71

38. Rubinfeld, B., Souza, B., Albert, I., Munemitsu, S. and **Polakis P.** 1995 The APC protein and E-cadherin form similar but independent complexes with α-catenin, β-catenin and Plakoglobin. **J. Biol. Chem.** 270, 5549-5555

39. Munemitsu, S., Albert, I., Souza, B., Rubinfeld, B., and **Polakis, P.** 1995 Regulation of intracellular β-catenin levels by the APC tumor suppressor gene. **Proc. Natl. Acad. Sci.** 92, 3046-3050.

40. Lock, P., Fumagalli, S., **Polakis, P.** McCormick, F. and Courtneidge, S. A. 1996 The human p62 cDNA encodes Sam68 and not the rasGAP-associated p62 protein. **Cell** 84, 23-24.

41. Papkoff, J., Rubinfeld, B., Schryver, B. and **Polakis, P.** 1996 Wnt-1 regulates free pools of catenins and stabilizes APC-catenin complexes. **Mol. Cell. Biol.** 16, 2128-2134.

42. Rubinfeld, B., Albert, I., Porfiri, E., Fiol, C., Munemitsu, S. and **Polakis, P.** 1996 Binding of GSK3β to the APC-β-catenin complex and regulation of complex assembly. **Science** 272, 1023-1026.

43. Munemitsu, S., Albert, I., Rubinfeld, B. and **Polakis, P.** 1996 Deletion of amino-terminal structure stabilizes β-catenin in vivo and promotes the hyperphosphorylation of the APC tumor suppressor protein. **Mol. Cell. Biol.**16, 4088-4094.

44. Hart, M. J., Callow, M. G., Sousa, B. and **Polakis P.** 1996 IQGAP1, a calmodulin binding protein witha rasGAP related domain, is a potential effector for cdc42Hs. **EMBO J.** 15, 2997-3005.

45. Nathke, I. S., Adams, C. L., **Polakis, P.**, Sellin, J. and Nelson, W. J. 1996 The adenomatous polyposis coli (APC) tumor suppressor protein is localized to plasma membrane sites involved in active epithelial cell migration. **J. Cell. Biol.** 134, 165-180.

46. Hart, M. J., Sharma, S., elMasry, N., Qui, R-G., McCabe, P., **Polakis, P.** and Bollag, G. 1996 Identification of a novel guanine nucleotide exchange factor for the rho GTPase. **J. Biol. Chem.** 271, 25452.

47. Thomas JE, Smith M, Rubinfeld B, Gutowski M, Beckmann RP, and **Polakis P.** 1996 Subcellular localization and analysis of apparent 180-kDa and 220-kDa proteins of the breast cancer susceptibility gene, BRCA1. **J. Biol. Chem.** 1996 271, 28630-28635

48. Hayashi, S., Rubinfeld, B., Souza, B., **Polakis, P.**, Wieschaus, E., and Levine, A. 1997 A Drosophila homolog of the tumor suppressor adenomatous polyposis coli

down-regulates β -catenin but its zygotic expression is not essential for the regulation of armadillo. **Proc. Natl. Acad. Sci.** 94, 242-247.

49. Vleminckx, K., Rubinfeld, B., **Polakis, P.** and Gumbiner, B. 1997 The APC tumor suppressor protein induces a new axis in Xenopus embryos. **J. Cell. Biol.** 136, 411-420.

50. Rubinfeld, B., Robbins, P., El-Gamil, M., Albert, I., Porfiri, P. and **Polakis, P.** 1997 Stabilization of β-catenin by genetic defects in melanoma cell lines. **Science** 275, 1790-1792.

51. Polakis, P. The adenomatous polyposis coli (APC) tumor suppressor. 1997 **Biochem. Biophys. Acta**, 1332, F127-F147.

52. Rubinfeld, B., Albert, I., Porfiri, E., Munemitsu, S., and **Polakis, P** 1997 Loss of β-catenin regulation by the APC tumor suppressor protein correlates with loss of structure due to common somatic mutations of the gene. **Cancer Res.** 57, 4624-4630.

53. Porfiri, E., Rubinfeld, B., Albert, I., Hovanes. K., Waterman, M., and **Polakis, P.** 1997 Induction of a β-catenin-LEF-1 complex by wnt-1 and transforming mutants of β-catenin. **Oncogene** 15, 2833-2839.

54. Thomas JE, Smith M, Tonkinson JL, Rubinfeld B, and **Polakis P.**, 1997 Induction of phosphorylation on BRCA1 during the cell cycle and after DNA damage. **Cell Growth Differ.** 8, 801-809.

55. Hart, M., de los Santos, R., Albert, I., Rubinfeld, B., and **Polakis P.**, 1998 Down regulation of β-catenin by human Axin and its association with the adenomatous polyposis coli (APC) tumor suppressor, β-catenin and glycogen synthase kinase 3β. **Current Biology 8,** 573-581.

56. Polakis, P. 1998 The oncogenic activation of β-catenin. **Current Opinions in Genetics and Development** 9, 15-21

57. Matt Hart, Jean-Paul Concordet, Irina Lassot, Iris Albert, Rico del los Santos, Herve Durand, Christine Perret, Bonnee Rubinfled, Florence Margottin, Richard Benarous and **Paul Polakis.** 1999 The F-box protein β–TrCP associates with phosphorylated β–catenin and regulates its activity in the cell. **Current Biology** 9, 207-10.

58. Howard C. Crawford, Barbara M. Fingleton, Bonnee Rubinfeld, **Paul Polakis** and Lynn M. Matrisian 1999 The metalloproteinase matrilysin is a target of β–catenin transactivation in intestinal tumours. **Oncogene** 18, 2883-91.

59. Meng J, Glick JL, **Polakis P**, Casey PJ. 1999 Functional interaction between Galpha(z) and Rap1GAP suggests a novel form of cellular cross-talk. **J Biol Chem.** 17, 36663-9

60. Vijayasurian Easwaran, Virginia Song, **Paul Polakis** and Steve Byers 1999 The ubiquitin-proteosome pathway and serine kinase activity modulate APC mediated regulation of β-catenin-LEF signaling. **J. Biol. Chem**. 274(23):16641-5.

61 Polakis P, Hart M and Rubinfeld B. 1999 Defects in the regulation of beta-catenin
in colorectal cancer. **Adv Exp Med Biol**. 470, 23-32

62 Shen Z, Batzer A, Koehler JA, **Polakis P**, Schlessinger J, Lydon NB, Moran MF. 1999 Evidence for SH3 domain directed binding and phosphorylation of Sam68 by Src. **Oncogene**. 18, 4647-53

64. Thomas GM, Frame S, Goedert M, Nathke I, **Polakis P**, Cohen P. 1999 A GSK3- binding peptide from FRAT1 selectively inhibits the GSK3-catalysed phosphorylation of axin and beta-catenin. **FEBS Lett**. 458, 247-51.

65. Peifer M, **Polakis P**. 2000 Wnt signaling in oncogenesis and embryogenesis--a look outside the nucleus. **Science** 287,1606-9.

66. **Polakis P**. 2000 Wnt signaling and cancer. **Genes Dev**;14, 1837-1851.

67. Spink KE, **Polakis P**, Weis WI 2000 Structural basis of the Axin-adenomatous polyposis coli interaction. **EMBO J** 19, 2270-2279.

68. Szeto , W., Jiang, W., Tice, D.A., Rubinfeld, B., Hollingshead, P.G., Fong, S.E., Dugger, D.L., Pham, T., Yansura, D.E., Wong, T.A., Grimaldi, J.C., Corpuz, R.T., Singh J.S., Frantz, G.D., Devaux, B., Crowley, C.W., Schwall, R.H., Eberhard, D.A.,
Rastelli, L., **Polakis, P**. and Pennica, D. 2001 Overexpression of the Retinoic Acid-
Responsive Gene Stra6 in Human Cancers and its Synergistic Induction by Wnt-1 and
Retinoic Acid. **Cancer Res** 61, 4197-4204.

69. Rubinfeld B, Tice DA, **Polakis P**. 2001 Axin dependent phosphorylation of the adenomatous polyposis coli protein mediated by casein kinase 1 epsilon. **J Biol Chem**
276, 39037-39045.

70. **Polakis P**. 2001 More than one way to skin a catenin. **Cell** 2001 105, 563-566.

71. Tice DA, Soloviev I, **Polakis P**. 2002 Activation of the Wnt Pathway Interferes withSerum Response Element-driven Transcription of Immediate Early Genes. **J Biol.**
Chem. 277, 6118-6123.

72. Tice DA, Szeto W, Soloviev I, Rubinfeld B, Fong SE, Dugger DL, Winer J,

Williams PM, Wieand D, Smith V, Schwall RH, Pennnica D, **Polakis P**. 2002 Synergistic activation of tumor antigens by wnt-1 signaling and retinoic acid revealed by gene expression profiling. **J Biol Chem**. 277,14329-14335.

73. **Polakis, P**. 2002 Casein kinase I: A wnt'er of disconnect. **Curr. Biol**. 12, R499.

74. Mao,W. , Luis, E., Ross, S., Silva, J., Tan, C., Crowley, C., Chui, C., Franz, G., Senter, P., Koeppen, H., **Polakis, P.** 2004 EphB2 as a therapeutic antibody drug target for the treatment of colorectal cancer. **Cancer Res**. 64, 781-788.

75. Shibamoto, S., Winer, J., Williams, M., Polakis, P. 2003 A Blockade in Wnt signaling is activated following the differentiation of F9 teratocarcinoma cells. **Exp. Cell Res.** 29211-20.

76. Zhang Y, Eberhard DA, Frantz GD, Dowd P, Wu TD, Zhou Y, Watanabe C, Luoh SM, **Polakis P**, Hillan KJ, Wood WI, Zhang Z. 2004 GEPIS--quantitative gene expression profiling in normal and cancer tissues. **Bioinformatics**, April 8

SECOND DECLARATION OF PAUL POLAKIS, Ph.D.

I, Paul Polakis, Ph.D., declare and say as follows:

1. I am currently employed by Genentech, Inc. where my job title is Staff Scientist.

2. Since joining Genentech in 1999, one of my primary responsibilities has been leading Genentech's Tumor Antigen Project, which is a large research project with a primary focus on identifying tumor cell markers that find use as targets for both the diagnosis and treatment of cancer in humans.

3. As I stated in my previous Declaration dated May 7, 2004 (attached as Exhibit A), my laboratory has been employing a variety of techniques, including microarray analysis, to identify genes which are differentially expressed in human tumor tissue relative to normal human tissue. The primary purpose of this research is to identify proteins that are abundantly expressed on certain human tumor tissue(s) and that are either (i) not expressed, or (ii) expressed at detectably lower levels, on normal tissue(s).

4. In the course of our research using microarray analysis, we have identified approximately 200 gene transcripts that are present in human tumor tissue at significantly higher levels than in normal human tissue. To date, we have successfully generated antibodies that bind to 31 of the tumor antigen proteins expressed from these differentially expressed gene transcripts and have used these antibodies to quantitatively determine the level of production of these tumor antigen proteins in both human tumor tissue and normal tissue. We have then quantitatively compared the levels of mRNA and protein in both the tumor and normal tissues analyzed. The results of these analyses are attached herewith as Exhibit B. In Exhibit B, "+" means that the mRNA or protein was detectably overexpressed in the tumor tissue relative to normal tissue and "-" means that no detectable overexpression was observed in the tumor tissue relative to normal tissue.

5. As shown in Exhibit B, of the 31 genes identified as being detectably overexpressed in human tumor tissue as compared to normal human tissue at the mRNA level, 28 of them (i.e., greater than 90%) are also detectably overexpressed in human tumor tissue as compared to normal human tissue at the protein level. As such, in the cases where we have been able to quantitatively measure both (i) mRNA and (ii) protein levels in both (i) tumor tissue and (ii) normal tissue, we have observed that in the vast majority of cases, there is a very strong correlation between increases in mRNA expression and increases in the level of protein encoded by that mRNA.

6. Based upon my own experience accumulated in more than 20 years of research, including the data discussed in paragraphs 4-5 above and my knowledge of the relevant scientific literature, it is my considered scientific opinion that for human genes, an increased level of mRNA in a tumor tissue relative to a normal tissue more often than not correlates to a similar increase in abundance of the encoded protein in the tumor tissue relative to the normal tissue. In fact, it remains a generally accepted working assumption in molecular biology that increased mRNA levels are more often than not predictive of elevated levels of the encoded protein. In fact, an entire industry focusing on the research and development of therapeutic antibodies to treat a variety of human diseases, such as cancer, operates on this working assumption.

7. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information or belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful statements may jeopardize the validity of the application or any patent issued thereon.

Dated: 3-29-06

By: Paul Polakis

Paul Polakis, Ph.D.

DECLARATION OF PAUL POLAKIS, Ph.D.

I, Paul Polakis, Ph.D., declare and say as follows:

1. I was awarded a Ph.D. by the Department of Biochemistry of the Michigan State University in 1984. My scientific Curriculum Vitae is attached to and forms part of this Declaration (Exhibit A).

2. I am currently employed by Genentech, Inc. where my job title is Staff Scientist. Since joining Genentech in 1999, one of my primary responsibilities has been leading Genentech's Tumor Antigen Project, which is a large research project with a primary focus on identifying tumor cell markers that find use as targets for both the diagnosis and treatment of cancer in humans.

3. As part of the Tumor Antigen Project, my laboratory has been analyzing differential expression of various genes in tumor cells relative to normal cells. The purpose of this research is to identify proteins that are abundantly expressed on certain tumor cells and that are either (i) not expressed, or (ii) expressed at lower levels, on corresponding normal cells. We call such differentially expressed proteins "tumor antigen proteins". When such a tumor antigen protein is identified, one can produce an antibody that recognizes and binds to that protein. Such an antibody finds use in the diagnosis of human cancer and may ultimately serve as an effective therapeutic in the treatment of human cancer.

4. In the course of the research conducted by Genentech's Tumor Antigen Project, we have employed a variety of scientific techniques for detecting and studying differential gene expression in human tumor cells relative to normal cells, at genomic DNA, mRNA and protein levels. An important example of one such technique is the well known and widely used technique of microarray analysis which has proven to be extremely useful for the identification of mRNA molecules that are differentially expressed in one tissue or cell type relative to another. In the course of our research using microarray analysis, we have identified approximately 200 gene transcripts that are present in human tumor cells at significantly higher levels than in corresponding normal human cells. To date, we have generated antibodies that bind to about 30 of the tumor antigen proteins expressed from these differentially expressed gene transcripts and have used these antibodies to quantitatively determine the level of production of these tumor antigen proteins in both human cancer cells and corresponding normal cells. We have then compared the levels of mRNA and protein in both the tumor and normal cells analyzed.

5. From the mRNA and protein expression analyses described in paragraph 4 above, we have observed that there is a strong correlation between changes in the level of mRNA present in any particular cell type and the level of protein

expressed from that mRNA in that cell type. In approximately 80% of our observations we have found that increases in the level of a particular mRNA correlates with changes in the level of protein expressed from that mRNA when human tumor cells are compared with their corresponding normal cells.

6. Based upon my own experience accumulated in more than 20 years of research, including the data discussed in paragraphs 4 and 5 above and my knowledge of the relevant scientific literature, it is my considered scientific opinion that for human genes, an increased level of mRNA in a tumor cell relative to a normal cell typically correlates to a similar increase in abundance of the encoded protein in the tumor cell relative to the normal cell. In fact, it remains a central dogma in molecular biology that increased mRNA levels are predictive of corresponding increased levels of the encoded protein. While there have been published reports of genes for which such a correlation does not exist, it is my opinion that such reports are exceptions to the commonly understood general rule that increased mRNA levels are predictive of corresponding increased levels of the encoded protein.

7. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information or belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful statements may jeopardize the validity of the application or any patent issued thereon.

Dated: 5/07/04

By: Paul Polakis

Paul Polakis, Ph.D.

SV 2031808 v1

CURRICULUM VITAE

PAUL G. POLAKIS
Staff Scientist
Genentech, Inc
1 DNA Way, MS#40
S. San Francisco, CA 94080

EDUCATION:

Ph.D., Biochemistry, Department of Biochemistry,
Michigan State University (1984)

B.S., Biology. College of Natural Science, Michigan State University (1977)

PROFESSIONAL EXPERIENCE:

2002-present	Staff Scientist, Genentech, Inc S. San Francisco, CA
1999- 2002	Senior Scientist, Genentech, Inc., S. San Francisco, CA
1997 -1999	Research Director Onyx Pharmaceuticals, Richmond, CA
1992- 1996	Senior Scientist, Project Leader, Onyx Pharmaceuticals, Richmond, CA
1991-1992	Senior Scientist, Chiron Corporation, Emeryville, CA.
1989-1991	Scientist, Cetus Corporation, Emeryville CA.
1987-1989	Postdoctoral Research Associate, Genentech, Inc., South SanFrancisco, CA.
1985-1987	Postdoctoral Research Associate, Department of Medicine, Duke University Medical Center, Durham, NC

1984-1985	Assistant Professor, Department of Chemistry, Oberlin College, Oberlin, Ohio
1980-1984	Graduate Research Assistant, Department of Biochemistry, Michigan State University East Lansing, Michigan

PUBLICATIONS:

1. **Polakis, P G.** and Wilson, J. E. 1982 Purification of a Highly Bindable Rat Brain Hexokinase by High Performance Liquid Chromatography. **Biochem. Biophys. Res. Commun**. 107, 937-943.

2. **Polakis, P.G.** and Wilson, J. E. 1984 Proteolytic Dissection of Rat Brain Hexokinase: Determination of the Cleavage Pattern during Limited Digestion with Trypsin. **Arch. Biochem. Biophys**. 234, 341-352.

3. **Polakis, P. G.** and Wilson, J. E. 1985 An Intact Hydrophobic N-Terminal Sequence is Required for the Binding Rat Brain Hexokinase to Mitochondria. **Arch. Biochem. Biophys**. 236, 328-337.

4. Uhing, R.J., **Polakis,P.G.** and Snyderman, R. 1987 Isolaton of GTP-binding Proteins from Myeloid HL60 Cells. **J. Biol. Chem**. 262, 15575-15579.

5. **Polakis, P.G.**, Uhing, R.J. and Snyderman, R. 1988 The Formylpeptide Chemoattractant Receptor Copurifies with a GTP-binding Protein Containing a Distinct 40 kDa Pertussis Toxin Substrate. **J. Biol. Chem**. 263, 4969-4979.

6. Uhing, R. J., Dillon, S., **Polakis, P. G.**, Truett, A. P. and Snyderman, R. 1988 Chemoattractant Receptors and Signal Transduction Processes in Cellular and Molecular Aspects of Inflammation (Poste, G. and Crooke, S. T. eds.) pp 335-379.

7. **Polakis, P.G.**, Evans, T. and Snyderman 1989 Multiple Chromatographic Forms of the Formylpeptide Chemoattractant Receptor and their Relationship to GTP-binding Proteins. **Biochem. Biophys. Res. Commun**. 161, 276-283.

8. **Polakis, P. G.**, Snyderman, R. and Evans, T. 1989 Characterization of G25K, a GTP-binding Protein Containing a Novel Putative Nucleotide Binding Domain. **Biochem. Biophys. Res. Comun.** 160, 25-32.

9. **Polakis, P.**, Weber,R.F., Nevins,B., Didsbury, J. Evans,T. and Snyderman, R. 1989 Identification of the ral and rac1 Gene Products, Low Molecular Mass GTP-binding Proteins from Human Platelets. **J. Biol. Chem**. 264, 16383-16389.

10. Snyderman, R., Perianin, A., Evans, T., **Polakis, P.** and Didsbury, J. 1989 G Proteins and Neutrophil Function. In ADP-Ribosylating Toxins and G Proteins: Insights into Signal Transduction. (J. Moss and M. Vaughn, eds.) Amer. Soc. Microbiol. pp. 295-323.

11. Hart, M.J., **Polakis, P.G.**, Evans, T. and Cerrione, R.A. 1990 The Identification and Charaterization of an Epidermal Growth Factor-Stimulated Phosphorylation of a Specific Low Molecular Mass GTP-binding Protein in a Reconstituted Phospholipid Vesicle System. **J. Biol. Chem**. 265, 5990-6001.

12. Yatani, A., Okabe, K., **Polakis, P.** Halenbeck, R. McCormick, F. and Brown, A. M. 1990 ras p21 and GAP Inhibit Coupling of Muscarinic Receptors to Atrial K^+ Channels. **Cell**. 61, 769-776.

13. Munemitsu, S., Innis, M.A., Clark, R., McCormick, F., Ullrich, A. and **Polakis, P.G.** 1990 Molecular Cloning and Expression of a G25K cDNA, the Human Homolog of the Yeast Cell Cycle Gene CDC42. **Mol. Cell. Biol**. 10, 5977-5982.

14. **Polakis, P.G.** Rubinfeld, B. Evans, T. and McCormick, F. 1991 Purification of Plasma Membrane-Associated GTPase Activating Protein (GAP) Specific for rap-1/krev-1 from HL60 Cells. **Proc. Natl. Acad. Sci. USA** 88, 239-243.

15. Moran, M. F., **Polakis, P.**, McCormick, F., Pawson, T. and Ellis, C. 1991 Protein Tyrosine Kinases Regulate the Phosphorylation, Protein Interactions, Subcellular Distribution, and Activity of p21ras GTPase Activating Protein. **Mol. Cell. Biol.** 11, 1804-1812

16. Rubinfeld, B., Wong, G., Bekesi, E. Wood, A. McCormick, F. and **Polakis, P. G.** 1991 A Synthetic Peptide Corresponding to a Sequence in the GTPase Activating Protein Inhibits $p21^{ras}$ Stimulation and Promotes Guanine Nucleotide Exchange. **Internatl. J. Peptide and Prot. Res**. 38, 47-53.

17. Rubinfeld, B., Munemitsu, S., Clark, R., Conroy, L., Watt, K., Crosier, W., McCormick, F., and **Polakis, P.** 1991 Molecular Cloning of a GTPase Activating Protein Specific for the Krev-1 Protein $p21^{rap1}$. **Cell** 65, 1033-1042.

18. Zhang, K. Papageorge, A., G., Martin, P., Vass, W. C., Olah, Z., **Polakis, P.**, McCormick, F. and Lowy, D, R. 1991 Heterogenous Amino Acids in RAS and Rap1A Specifying Sensitivity to GAP Proteins. **Science** 254, 1630-1634.

19. Martin, G., Yatani, A., Clark, R., **Polakis, P.**, Brown, A. M. and McCormick, F. 1992 GAP Domains Responsible for $p21^{ras}$-dependent Inhibition of Muscarinic Atrial K^+ Channel Currents. **Science** 255, 192-194.

20. McCormick, F., Martin, G. A., Clark, R., Bollag, G. and **Polakis, P** . 1992 Regulation of p21ras by GTPase Activating Proteins. Cold Spring Harbor **Symposia on Quantitative Biology**. Vol. 56, 237-241.

21. Pronk, G. B., **Polakis, P.**, Wong, G., deVries-Smits, A. M., Bos J. L. and McCormick, F. 1992 $p60^{v\text{-}src}$ Can Associate with and Phosphorylate the $p21^{ras}$ GTPase Activating Protein. **Oncogene** 7,389-394.

22. **Polakis P.** and McCormick, F. 1992 Interactions Between $p21^{ras}$ Proteins and Their GTPase Activating Proteins. In **Cancer Surveys** (Franks, L. M., ed.) 12, 25-42.

23. Wong, G., Muller, O., Clark, R., Conroy, L., Moran, M., **Polakis,P.** and McCormick, F. 1992 Molecular coloning and nucleic acid binding properties of the GAP-associated tyrosine phosphoprotein p62. **Cell** 69, 551-558.

24. Polakis, P., Rubinfeld, B. and McCormick, F. 1992 Phosphorylation of rap1GAP in vivo and by cAMP-dependent Kinase and the Cell Cycle $p34^{cdc2}$ Kinase in vitro. **J. Biol. Chem**. 267, 10780-10785.

25. McCabe, P.C., Haubrauck, H., **Polakis, P.**, McCormick, F., and Innis, M. A. 1992 Functional Interactions Between $p21^{rap1A}$ and Components of the Budding pathway of *Saccharomyces cerevisiae*. **Mol. Cell. Biol**. 12, 4084-4092.

26. Rubinfeld, B., Crosier, W.J., Albert, I., Conroy,L., Clark, R., McCormick, F. and **Polakis, P**. 1992 Localization of the rap1GAP Catalytic Domain and Sites of Phosphorylation by Mutational Analysis. **Mol. Cell . Biol**. 12, 4634-4642.

27. Ando, S., Kaibuchi, K., Sasaki, K., Hiraoka, T., Nishiyama, T., Mizuno, T., Asada, M., Nunoi, H., Matsuda, I., Matsuura, Y., **Polakis, P.**, McCormick, F. and Takai, Y. 1992 Post-translational processing of rac p21s is important both for their interaction with the GDP/GTP exchange proteins and for their activation of NADPH oxidase. **J. Biol. Chem**. 267, 25709-25713.

28. Janoueix-Lerosey, I., **Polakis, P.**, Tavitian, A. and deGunzberg, J. 1992 Regulation of the GTPase activity of the ras-related rap2 protein. **Biochem. Biophys. Res. Commun**. 189, 455-464.

29. Polakis, P. 1993 GAPs Specific for the rap1/Krev-1 Protein. in GTP-binding Proteins: the *ras*-superfamily. (J.C. LaCale and F. McCormick, eds.) 445-452.

30. Polakis, P. and McCormick, F. 1993 Structural requirements for the interaction of $p21^{ras}$ with GAP, exchange factors, and its bological effector target. **J. Biol Chem**. 268, 9157-9160.

31. Rubinfeld, B., Souza, B. Albert, I., Muller, O., Chamberlain, S., Masiarz, F., Munemitsu, S. and **Polakis, P.** 1993 Association of the APC gene product with beta- catenin. **Science** 262, 1731-1734.

32. Weiss, J., Rubinfeld, B., **Polakis, P.**, McCormick, F. Cavenee, W. A. and Arden, K. 1993 The gene for human rap1-GTPase activating protein (rap1GAP) maps to chromosome 1p35-1p36.1. **Cytogenet. Cell Genet**. 66, 18-21.

33. Sato, K. Y., **Polakis, P.**, Haubruck, H., Fasching, C. L., McCormick, F. and Stanbridge, E. J. 1994 Analysis of the tumor suppressor acitvity of the K-rev gene in human tumor cell lines. **Cancer Res**. 54, 552-559.

34. Janoueix-Lerosey, I., Fontenay, M., Tobelem, G., Tavitian, A., **Polakis, P.** and DeGunzburg, J. 1994 Phosphorylation of rap1GAP during the cell cycle. **Biochem. Biophys. Res. Commun**. 202, 967-975

35. Munemitsu, S., Souza, B., Mueller, O., Albert, I., Rubinfeld, B., and **Polakis, P.** 1994 The APC gene product associates with microtubules in vivo and affects their assembly in vitro. **Cancer Res**. 54, 3676-3681.

36. Rubinfeld, B. and **Polakis, P.** 1995 Purification of baculovirus produced rap1GAP. **Methods Enz.** 255,31

37. Polakis, P. 1995 Mutations in the APC gene and their implications for protein structure and function. **Current Opinions in Genetics and Development** 5, 66-71

38. Rubinfeld, B., Souza, B., Albert, I., Munemitsu, S. and **Polakis P.** 1995 The APC protein and E-cadherin form similar but independent complexes with α-catenin, β-catenin and Plakoglobin. **J. Biol. Chem.** 270, 5549-5555

39. Munemitsu, S., Albert, I., Souza, B., Rubinfeld, B., and **Polakis, P.** 1995 Regulation of intracellular β-catenin levels by the APC tumor suppressor gene. **Proc. Natl. Acad. Sci.** 92, 3046-3050.

40. Lock, P., Fumagalli, S., **Polakis, P.** McCormick, F. and Courtneidge, S. A. 1996 The human p62 cDNA encodes Sam68 and not the rasGAP-associated p62 protein. **Cell** 84, 23-24.

41. Papkoff, J., Rubinfeld, B., Schryver, B. and **Polakis, P.** 1996 Wnt-1 regulates free pools of catenins and stabilizes APC-catenin complexes. **Mol. Cell. Biol.** 16, 2128-2134.

42. Rubinfeld, B., Albert, I., Porfiri, E., Fiol, C., Munemitsu, S. and **Polakis, P.** 1996 Binding of GSK3β to the APC-β-catenin complex and regulation of complex assembly. **Science** 272, 1023-1026.

43. Munemitsu, S., Albert, I., Rubinfeld, B. and **Polakis, P.** 1996 Deletion of amino-terminal structure stabilizes β-catenin in vivo and promotes the hyperphosphorylation of the APC tumor suppressor protein. **Mol. Cell. Biol.**16, 4088-4094.

44. Hart, M. J., Callow, M. G., Sousa, B. and **Polakis P.** 1996 IQGAP1, a calmodulin binding protein witha rasGAP related domain, is a potential effector for cdc42Hs. **EMBO J.** 15, 2997-3005.

45. Nathke, I. S., Adams, C. L., **Polakis, P.**, Sellin, J. and Nelson, W. J. 1996 The adenomatous polyposis coli (APC) tumor suppressor protein is localized to plasma membrane sites involved in active epithelial cell migration. **J. Cell. Biol.** 134, 165-180.

46. Hart, M. J., Sharma, S., elMasry, N., Qui, R-G., McCabe, P., **Polakis, P.** and Bollag, G. 1996 Identification of a novel guanine nucleotide exchange factor for the rho GTPase. **J. Biol. Chem.** 271, 25452.

47. Thomas JE, Smith M, Rubinfeld B, Gutowski M, Beckmann RP, and **Polakis P.** 1996 Subcellular localization and analysis of apparent 180-kDa and 220-kDa proteins of the breast cancer susceptibility gene, BRCA1. **J. Biol. Chem.** 1996 271, 28630-28635

48. Hayashi, S., Rubinfeld, B., Souza, B., **Polakis, P.**, Wieschaus, E., and Levine, A. 1997 A Drosophila homolog of the tumor suppressor adenomatous polyposis coli

down-regulates β -catenin but its zygotic expression is not essential for the regulation of armadillo. **Proc. Natl. Acad. Sci.** 94, 242-247.

49. Vleminckx, K., Rubinfeld, B., **Polakis, P.** and Gumbiner, B. 1997 The APC tumor suppressor protein induces a new axis in Xenopus embryos. **J. Cell. Biol.** 136, 411-420.

50. Rubinfeld, B., Robbins, P., El-Gamil, M., Albert, I., Porfiri, P. and **Polakis, P.** 1997 Stabilization of β-catenin by genetic defects in melanoma cell lines. **Science** 275, 1790-1792.

51. Polakis, P. The adenomatous polyposis coli (APC) tumor suppressor. 1997 **Biochem. Biophys. Acta**, 1332, F127-F147.

52. Rubinfeld, B., Albert, I., Porfiri, E., Munemitsu, S., and **Polakis, P** 1997 Loss of β-catenin regulation by the APC tumor suppressor protein correlates with loss of structure due to common somatic mutations of the gene. **Cancer Res.** 57, 4624-4630.

53. Porfiri, E., Rubinfeld, B., Albert, I., Hovanes. K., Waterman, M., and **Polakis, P.** 1997 Induction of a β-catenin-LEF-1 complex by wnt-1 and transforming mutants of β-catenin. **Oncogene** 15, 2833-2839.

54. Thomas JE, Smith M, Tonkinson JL, Rubinfeld B, and **Polakis P.,** 1997 Induction of phosphorylation on BRCA1 during the cell cycle and after DNA damage. **Cell Growth Differ.** 8, 801-809.

55. Hart, M., de los Santos, R., Albert, I., Rubinfeld, B., and **Polakis P.**, 1998 Down regulation of β-catenin by human Axin and its association with the adenomatous polyposis coli (APC) tumor suppressor, β-catenin and glycogen synthase kinase 3β. **Current Biology 8,** 573-581.

56. Polakis, P. 1998 The oncogenic activation of β-catenin. **Current Opinions in Genetics and Development** 9, 15-21

57. Matt Hart, Jean-Paul Concordet, Irina Lassot, Iris Albert, Rico del los Santos, Herve Durand, Christine Perret, Bonnee Rubinfled, Florence Margottin, Richard Benarous and **Paul Polakis.** 1999 The F-box protein β–TrCP associates with phosphorylated β–catenin and regulates its activity in the cell. **Current Biology** 9, 207-10.

58. Howard C. Crawford, Barbara M. Fingleton, Bonnee Rubinfeld, **Paul Polakis** and Lynn M. Matrisian 1999 The metalloproteinase matrilysin is a target of β–catenin transactivation in intestinal tumours. **Oncogene** 18, 2883-91.

59. Meng J, Glick JL, **Polakis P**, Casey PJ. 1999 Functional interaction between Galpha(z) and Rap1GAP suggests a novel form of cellular cross-talk. **J Biol Chem.** 17, 36663-9

60. Vijayasurian Easwaran, Virginia Song, **Paul Polakis** and Steve Byers 1999 The ubiquitin-proteosome pathway and serine kinase activity modulate APC mediated regulation of β-catenin-LEF signaling. **J. Biol. Chem.** 274(23):16641-5.

61 Polakis P, Hart M and Rubinfeld B. 1999 Defects in the regulation of beta-catenin
in colorectal cancer. **Adv Exp Med Biol.** 470, 23-32

62 Shen Z, Batzer A, Koehler JA, **Polakis P**, Schlessinger J, Lydon NB, Moran MF. 1999 Evidence for SH3 domain directed binding and phosphorylation of Sam68 by Src. **Oncogene**. 18, 4647-53

64. Thomas GM, Frame S, Goedert M, Nathke I, **Polakis P**, Cohen P. 1999 A GSK3- binding peptide from FRAT1 selectively inhibits the GSK3-catalysed phosphorylation of axin and beta-catenin. **FEBS Lett**. 458, 247-51.

65. Peifer M, **Polakis P**. 2000 Wnt signaling in oncogenesis and embryogenesis--a look outside the nucleus. **Science** 287,1606-9.

66. **Polakis P.** 2000 Wnt signaling and cancer. **Genes Dev**;14, 1837-1851.

67. Spink KE, **Polakis P**, Weis WI 2000 Structural basis of the Axin-adenomatous polyposis coli interaction. **EMBO J** 19, 2270-2279.

68. Szeto , W., Jiang, W., Tice, D.A., Rubinfeld, B., Hollingshead, P.G., Fong, S.E., Dugger, D.L., Pham, T., Yansura, D.E., Wong, T.A., Grimaldi, J.C., Corpuz, R.T., Singh J.S., Frantz, G.D., Devaux, B., Crowley, C.W., Schwall, R.H., Eberhard, D.A.,
Rastelli, L., **Polakis, P.** and Pennica, D. 2001 Overexpression of the Retinoic Acid-
Responsive Gene Stra6 in Human Cancers and its Synergistic Induction by Wnt-1 and
Retinoic Acid. **Cancer Res** 61, 4197-4204.

69. Rubinfeld B, Tice DA, **Polakis P**. 2001 Axin dependent phosphorylation of the adenomatous polyposis coli protein mediated by casein kinase 1 epsilon. **J Biol Chem**
276, 39037-39045.

70. **Polakis P**. 2001 More than one way to skin a catenin. **Cell** 2001 105, 563-566.

71. Tice DA, Soloviev I, **Polakis P**. 2002 Activation of the Wnt Pathway Interferes withSerum Response Element-driven Transcription of Immediate Early Genes. **J Biol.**
Chem. 277, 6118-6123.

72. Tice DA, Szeto W, Soloviev I, Rubinfeld B, Fong SE, Dugger DL, Winer J,

Williams PM, Wieand D, Smith V, Schwall RH, Pennnica D, **Polakis P.** 2002 Synergistic activation of tumor antigens by wnt-1 signaling and retinoic acid revealed by gene expression profiling. **J Biol Chem.** 277,14329-14335.

73. **Polakis, P.** 2002 Casein kinase I: A wnt'er of disconnect. **Curr. Biol.** 12, R499.

74. Mao,W. , Luis, E., Ross, S., Silva, J., Tan, C., Crowley, C., Chui, C., Franz, G., Senter, P., Koeppen, H., **Polakis, P.** 2004 EphB2 as a therapeutic antibody drug target for the treatment of colorectal cancer. **Cancer Res.** 64, 781-788.

75. Shibamoto, S., Winer, J., Williams, M., Polakis, P. 2003 A Blockade in Wnt signaling is activated following the differentiation of F9 teratocarcinoma cells. **Exp. Cell Res.** 29211-20.

76. Zhang Y, Eberhard DA, Frantz GD, Dowd P, Wu TD, Zhou Y, Watanabe C, Luoh SM, **Polakis P,** Hillan KJ, Wood WI, Zhang Z. 2004 GEPIS--quantitative gene expression profiling in normal and cancer tissues. **Bioinformatics**, April 8

EXHIBIT B

	tumor mRNA	tumor IHC
UNQ2525	+	+
UNQ2378	+	+
UNQ972	+	-
UNQ97671	+	+
UNQ2964	+	+
UNQ323	+	+
UNQ1655	+	+
UNQ2333	+	+
UNQ9638	+	+
UNQ8209	+	+
UNQ6507	+	+
UNQ8196	+	+
UNQ9109	+	+
UNQ100	+	+
UNQ178	+	+
UNQ1477	+	+
UNQ1839	+	+
UNQ2079	+	+
UNQ8782	+	+
UNQ9646	+	-
UNQ111	+	+
UNQ3079	+	+
UNQ8175	+	+
UNQ9509	+	+
UNQ10978	+	-
UNQ2103	+	+
UNQ1563	+	+
UNQ16188	+	+
UNQ13589	+	+
UNQ1078	+	+
UNQ879	+	+

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant	:	Ashkenazi et al.
App. No.	:	09/903,925
Filed	:	July 11, 2001
For	:	SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
Examiner	:	Hamud, Fozia M

Group Art Unit 1647

CERTIFICATE OF EXPRESS MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to Commissioner of Patents, Washington D.C. 20231 on:

30 May 2006
(Date)
C. Noel Kamen

Commissioner of Patents
P.O. Box 1450
Alexandria, VA 22313-1450

DECLARATION OF AVI ASHKENAZI, Ph.D UNDER 37 C.F.R. § 1.132

I, Avi Ashkenazi, Ph.D. declare and say as follows: -

1. I am Director and Staff Scientist at the Molecular Oncology Department of Genentech, Inc., South San Francisco, CA 94080.

2. I joined Genentech in 1988 as a postdoctoral fellow. Since then, I have investigated a variety of cellular signal transduction mechanisms, including apoptosis, and have developed technologies to modulate such mechanisms as a means of therapeutic intervention in cancer and autoimmune disease. I am currently involved in the investigation of a series of secreted proteins over-expressed in tumors, with the aim to identify useful targets for the development of therapeutic antibodies for cancer treatment.

3. My scientific Curriculum Vitae, including my list of publications, is attached to and forms part of this Declaration (Exhibit A).

4. Gene amplification is a process in which chromosomes undergo changes to contain multiple copies of certain genes that normally exist as a single copy, and is an important factor in the pathophysiology of cancer. Amplification of certain genes (e.g., Myc or Her2/Neu)

gives cancer cells a growth or survival advantage relative to normal cells, and might also provide a mechanism of tumor cell resistance to chemotherapy or radiotherapy.

5. If gene amplification results in over-expression of the mRNA and the corresponding gene product, then it identifies that gene product as a promising target for cancer therapy, for example by the therapeutic antibody approach. Even in the absence of over-expression of the gene product, amplification of a cancer marker gene - as detected, for example, by the reverse transcriptase TaqMan® PCR or the fluorescence *in situ* hybridization (FISH) assays -is useful in the diagnosis or classification of cancer, or in predicting or monitoring the efficacy of cancer therapy. An increase in gene copy number can result not only from intrachromosomal changes but also from chromosomal aneuploidy. It is important to understand that detection of gene amplification can be used for cancer diagnosis even if the determination includes measurement of chromosomal aneuploidy. Indeed, as long as a significant difference relative to normal tissue is detected, it is irrelevant if the signal originates from an increase in the number of gene copies per chromosome and/or an abnormal number of chromosomes.

6. I understand that according to the Patent Office, absent data demonstrating that the increased copy number of a gene in certain types of cancer leads to increased expression of its product, gene amplification data are insufficient to provide substantial utility or well established utility for the gene product (the encoded polypeptide), or an antibody specifically binding the encoded polypeptide. However, even when amplification of a cancer marker gene does not result in significant over-expression of the corresponding gene product, this very absence of gene product over-expression still provides significant information for cancer diagnosis and treatment. Thus, if over-expression of the gene product does not parallel gene amplification in certain tumor types but does so in others, then parallel monitoring of gene amplification and gene product over-expression enables more accurate tumor classification and hence better determination of suitable therapy. In addition, absence of over-expression is crucial information for the practicing clinician. If a gene is amplified but the corresponding gene product is not over-expressed, the clinician accordingly will decide not to treat a patient with agents that target that gene product.

7. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information or belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so

made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful statements may jeopardize the validity of the application or any patent issued thereon.

By: Avi Ashkenazi
Avi Ashkenazi, Ph.D.

Date: 9/15/03

SV 455281 v1
9/12/03 3:06 PM (39780.7000)

CURRICULUM VITAE

Avi Ashkenazi

July 2003

Personal:

Date of birth:	29 November, 1956
Address:	1456 Tarrytown Street, San Mateo, CA 94402
Phone:	(650) 578-9199 (home); (650) 225-1853 (office)
Fax:	(650) 225-6443 (office)
Email:	aa@gene.com

Education:

1983:	B.S. in Biochemistry, with honors, Hebrew University, Israel
1986:	Ph.D. in Biochemistry, Hebrew University, Israel

Employment:

1983-1986:	Teaching assistant, undergraduate level course in Biochemistry
1985-1986:	Teaching assistant, graduate level course on Signal Transduction
1986 - 1988:	Postdoctoral fellow, Hormone Research Dept., UCSF, and Developmental Biology Dept., Genentech, Inc., with J. Ramachandran
1988 - 1989:	Postdoctoral fellow, Molecular Biology Dept., Genentech, Inc., with D. Capon
1989 - 1993:	Scientist, Molecular Biology Dept., Genentech, Inc.
1994 -1996:	Senior Scientist, Molecular Oncology Dept., Genentech, Inc.
1996-1997:	Senior Scientist and Interim director, Molecular Oncology Dept., Genentech, Inc.
1997-1990:	Senior Scientist and preclinical project team leader, Genentech, Inc.
1999 -2002:	Staff Scientist in Molecular Oncology, Genentech, Inc.
2002-present:	Staff Scientist and Director in Molecular Oncology, Genentech, Inc.

Awards:

1988:	First prize, The Boehringer Ingelheim Award

Editorial:

Editorial Board Member: Current Biology
Associate Editor, Clinical Cancer Research.
Associate Editor, Cancer Biology and Therapy.

Refereed papers:

1. Gertler, A., Ashkenazi, A., and Madar, Z. Binding sites for human growth hormone and ovine and bovine prolactins in the mammary gland and liver of the lactating cow. *Mol. Cell. Endocrinol.* **34**, 51-57 (1984).
2. Gertler, A., Shamay, A., Cohen, N., Ashkenazi, A., Friesen, H., Levanon, A., Gorecki, M., Aviv, H., Hadari, D., and Vogel, T. Inhibition of lactogenic activities of ovine prolactin and human growth hormone (hGH) by a novel form of a modified recombinant hGH. *Endocrinology* **118**, 720-726 (1986).
3. Ashkenazi, A., Madar, Z., and Gertler, A. Partial purification and characterization of bovine mammary gland prolactin receptor. *Mol. Cell. Endocrinol.* **50**, 79-87 (1987).
4. Ashkenazi, A., Pines, M., and Gertler, A. Down-regulation of lactogenic hormone receptors in Nb2 lymphoma cells by cholera toxin. *Biochemistry Internatl.* **14**, 1065-1072 (1987).
5. Ashkenazi, A., Cohen, R., and Gertler, A. Characterization of lactogen receptors in lactogenic hormone-dependent and independent Nb2 lymphoma cell lines. *FEBS Lett.* **210**, 51-55 (1987).
6. Ashkenazi, A., Vogel, T., Barash, I., Hadari, D., Levanon, A., Gorecki, M., and Gertler, A. Comparative study on in vitro and in vivo modulation of lactogenic and somatotropic receptors by native human growth hormone and its modified recombinant analog. *Endocrinology* **121**, 414-419 (1987).
7. Peralta, E., Winslow, J., Peterson, G., Smith, D., Ashkenazi, A., Ramachandran, J., Schimerlik, M., and Capon, D. Primary structure and biochemical properties of an M2 muscarinic receptor. *Science* **236**, 600-605 (1987).
8. Peralta, E. Ashkenazi, A., Winslow, J., Smith, D., Ramachandran, J., and Capon, D. J. Distincnt primary structures, ligand-binding properties and tissue-specific expression of four human muscarinic acetylcholine receptors. *EMBO J.* **6**, 3923-3929 (1987).
9. Ashkenazi, A., Winslow, J., Peralta, E., Peterson, G., Schimerlik, M., Capon, D., and Ramachandran, J. An M2 muscarinic receptor subtype coupled to both adenylyl cyclase and phosphoinositide turnover. *Science* **238**, 672-675 (1987).

10. Pines, M., Ashkenazi, A., Cohen-Chapnik, N., Binder, L., and Gertler, A. Inhibition of the proliferation of Nb2 lymphoma cells by femtomolar concentrations of cholera toxin and partial reversal of the effect by 12-o-tetradecanoyl-phorbol-13-acetate. *J. Cell. Biochem.* **37**, 119-129 (1988).

11. Peralta, E. Ashkenazi, A., Winslow, J. Ramachandran, J., and Capon, D. Differential regulation of PI hydrolysis and adenylyl cyclase by muscarinic receptor subtypes. *Nature* **334**, 434-437 (1988).

12. Ashkenazi., A. Peralta, E., Winslow, J., Ramachandran, J., and Capon, D. Functionally distinct G proteins couple different receptors to PI hydrolysis in the same cell. *Cell* **56**, 487-493 (1989).

13. Ashkenazi, A., Ramachandran, J., and Capon, D. Acetylcholine analogue stimulates DNA synthesis in brain-derived cells via specific muscarinic acetylcholine receptor subtypes. *Nature* **340**, 146-150 (1989).

14. Lammare, D., Ashkenazi, A., Fleury, S., Smith, D., Sekaly, R., and Capon, D. The MHC-binding and gp120-binding domains of CD4 are distinct and separable. *Science* **245**, 743-745 (1989).

15. Ashkenazi., A., Presta, L., Marsters, S., Camerato, T., Rosenthal, K., Fendly, B., and Capon, D. Mapping the CD4 binding site for human immunodefficiency virus type 1 by alanine-scanning mutagenesis. *Proc. Natl. Acad. Sci. USA.* **87**, 7150-7154 (1990).

16. Chamow, S., Peers, D., Byrn, R., Mulkerrin, M., Harris, R., Wang, W., Bjorkman, P., Capon, D., and Ashkenazi, A. Enzymatic cleavage of a CD4 immunoadhesin generates crystallizable, biologically active Fd-like fragments. *Biochemistry* **29**, 9885-9891 (1990).

17. Ashkenazi, A., Smith, D., Marsters, S., Riddle, L., Gregory, T., Ho, D., and Capon, D. Resistance of primary isolates of human immunodefficiency virus type 1 to soluble CD4 is independent of CD4-rgp120 binding affinity. *Proc. Natl. Acad. Sci. USA.* **88**, 7056-7060 (1991).

18. Ashkenazi, A., Marsters, S., Capon, D., Chamow, S., Figari., I., Pennica, D., Goeddel., D., Palladino, M., and Smith, D. Protection against endotoxic shock by a tumor necrosis factor receptor immunoadhesin. *Proc. Natl. Acad. Sci. USA.* **88**, 10535-10539 (1991).

19. Moore, J., McKeating, J., Huang, Y., Ashkenazi, A., and Ho, D. Virions of primary HIV-1 isolates resistant to sCD4 neutralization differ in sCD4 affinity and glycoprotein gp120 retention from sCD4-sensitive isolates. *J. Virol.* **66**, 235-243 (1992).

20. Jin, H., Oksenberg, D., Ashkenazi, A., Peroutka, S., Duncan, A., Rozmahel., R., Yang, Y., Mengod, G., Palacios, J., and O'Dowd, B. Characterization of the human 5-hydroxytryptamine$_{1B}$ receptor. *J. Biol. Chem.* **267**, 5735-5738 (1992).

21. Marsters, A., Frutkin, A., Simpson, N., Fendly, B. and Ashkenazi, A. Identification of cysteine-rich domains of the type 1 tumor necrosis receptor involved in ligand binding. *J. Biol. Chem.* **267**, 5747-5750 (1992).

22. Chamow, S., Kogan, T., Peers, D., Hastings, R., Byrn, R., and Ashkenazi, A. Conjugation of sCD4 without loss of biological activity via a novel carbohydrate-directed cross-linking reagent. *J. Biol. Chem.* **267**, 15916-15922 (1992).

23. Oksenberg, D., Marsters, A., O'Dowd, B., Jin, H., Havlik, S., Peroutka, S., and Ashkenazi, A. A single amino-acid difference confers major pharmacologic variation between human and rodent 5-HT$_{1B}$ receptors. *Nature* **360**, 161-163 (1992).

24. Haak-Frendscho, M., Marsters, S., Chamow, S., Peers, D., Simpson, N., and Ashkenazi, A. Inhibition of interferon γ by an interferon γ receptor immunoadhesin. *Immunology* **79**, 594-599 (1993).

25. Penica, D., Lam, V., Weber, R., Kohr, W., Basa, L., Spellman, M., Ashkenazi, Shire, S., and Goeddel, D. Biochemical characterization of the extracellular domain of the 75-kd tumor necrosis factor receptor. *Biochemistry* **32**, 3131-3138. (1993).

26. Barfod, L., Zheng, Y., Kuang, W., Hart, M., Evans, T., Cerione, R., and Ashkenazi, A. Cloning and expression of a human CDC42 GTPase Activating Protein reveals a functional SH3-binding domain. *J. Biol. Chem.* **268**, 26059-26062 (1993).

27. Chamow, S., Zhang, D., Tan, X., Mhtre, S., Marsters, S., Peers, D., Byrn, R., Ashkenazi, A., and Yunghans, R. A humanized bispecific immunoadhesin-antibody that retargets CD3+ effectors to kill HIV-1-infected cells. *J. Immunol.* **153**, 4268-4280 (1994).

28. Means, R., Krantz, S., Luna, J., Marsters, S., and Ashkenazi, A. Inhibition of murine erythroid colony formation in vitro by iterferon γ and correction by interferon γ receptor immunoadhesin. *Blood* **83**, 911-915 (1994).

29. Haak-Frendscho, M., Marsters, S., Mordenti, J., Gillet, N., Chen, S., andAshkenazi, A. Inhibition of TNF by a TNF receptor immunoadhesin: comparison with an anti-TNF mAb. *J. Immunol.* **152**, 1347-1353 (1994).

30. Chamow, S., Kogan, T., Venuti, M., Gadek, T., Peers, D., Mordenti, J., Shak, S., and Ashkenazi, A. Modification of CD4 immunoadhesin with monomethoxy-PEG aldehyde via reductive alkilation. *Bioconj. Chem.* **5**, 133-140 (1994).

31. Jin, H., Yang, R., Marsters, S., Bunting, S., Wurm, F., Chamow, S., and Ashkenazi, A. Protection against rat endotoxic shock by p55 tumor necrosis factor (TNF) receptor immunoadhesin: comparison to anti-TNF monoclonal antibody. *J. Infect. Diseases* **170**, 1323-1326 (1994).

32. Beck, J., Marsters, S., Harris, R., Ashkenazi, A., and Chamow, S. Generation of soluble interleukin-1 receptor from an immunoadhesin by specific cleavage. *Mol. Immunol.* **31**, 1335-1344 (1994).

33. Pitti, B., Marsters, M., Haak-Frendscho, M., Osaka, G., Mordenti, J., Chamow, S., and Ashkenazi, A. Molecular and biological properties of an interleukin-1 receptor immunoadhesin. *Mol. Immunol.* **31**, 1345-1351 (1994).

34. Oksenberg, D., Havlik, S., Peroutka, S., and Ashkenazi, A. The third intracellular loop of the 5-HT2 receptor specifies effector coupling. *J. Neurochem.* **64**, 1440-1447 (1995).

35. Bach, E., Szabo, S., Dighe, A., Ashkenazi, A., Aguet, M., Murphy, K., and Schreiber, R. Ligand-induced autoregulation of IFN-γ receptor β chain expression in T helper cell subsets. *Science* **270**, 1215-1218 (1995).

36. Jin, H., Yang, R., Marsters, S., Ashkenazi, A., Bunting, S., Marra, M., Scott, R., and Baker, J. Protection against endotoxic shock by bactericidal/permeability-increasing protein in rats. *J. Clin. Invest.* **95**, 1947-1952 (1995).

37. Marsters, S., Penica, D., Bach, E., Schreiber, R., and Ashkenazi, A. Interferon γ signals via a high-affinity multisubunit receptor complex that contains two types of polypeptide chain. *Proc. Natl. Acad. Sci. USA.* **92**, 5401-5405 (1995).

38. Van Zee, K., Moldawer, L., Oldenburg, H., Thompson, W., Stackpole, S., Montegut, W., Rogy, M., Meschter, C., Gallati, H., Schiller, C., Richter, W., Loetcher, H., Ashkenazi, A., Chamow, S., Wurm, F., Calvano, S., Lowry, S., and Lesslauer, W. Protection against lethal E. coli bacteremia in baboons by pretreatment with a 55-kDa TNF receptor-Ig fusion protein, Ro45-2081. *J. Immunol.* **156**, 2221-2230 (1996).

39. Pitti, R., Marsters, S., Ruppert, S., Donahue, C., Moore, A., and Ashkenazi, A. Induction of apoptosis by Apo-2 Ligand, a new member of the tumor necrosis factor cytokine family. *J. Biol. Chem.* **271**, 12687-12690 (1996).

40. Marsters, S., Pitti, R., Donahue, C., Rupert, S., Bauer, K., and Ashkenazi, A. Activation of apoptosis by Apo-2 ligand is independent of FADD but blocked by CrmA. *Curr. Biol.* **6**, 1669-1676 (1996).

41. Marsters, S., Skubatch, M., Gray, C., and Ashkenazi, A. Herpesvirus entry mediator, a novel member of the tumor necrosis factor receptor family, activates the NF-κB and AP-1 transcription factors. *J. Biol. Chem.* **272**, 14029-14032 (1997).

42. Sheridan, J., Marsters, S., Pitti, R., Gurney, A., Skubatch, M., Baldwin, D., Ramakrishnan, L., Gray, C., Baker, K., Wood, W.I., Goddard, A., Godowski, P., and Ashkenazi, A. Control of TRAIL-induced apoptosis by a family of signaling and decoy receptors. *Science* **277**, 818-821 (1997).

43. Marsters, S., Sheridan, J., Pitti, R., Gurney, A., Skubatch, M., Balswin, D., Huang, A., Yuan, J., Goddard, A., Godowski, P., and Ashkenazi, A. A novel receptor for Apo2L/TRAIL contains a truncated death domain. *Curr. Biol.* **7**, 1003-1006 (1997).

44. Marsters, A., Sheridan, J., Pitti, R., Brush, J., Goddard, A., and Ashkenazi, A. Identification of a ligand for the death-domain-containing receptor Apo3. *Curr. Biol.* **8**, 525-528 (1998).

45. Rieger, J., Naumann, U., Glaser, T., Ashkenazi, A., and Weller, M. Apo2 ligand: a novel weapon against malignant glioma? *FEBS Lett.* **427**, 124-128 (1998).

46. Pender, S., Fell, J., Chamow, S., Ashkenazi, A., and MacDonald, T. A p55 TNF receptor immunoadhesin prevents T cell mediated intestinal injury by inhibiting matrix metalloproteinase production. *J. Immunol.* **160**, 4098-4103 (1998).

47. Pitti, R., Marsters, S., Lawrence, D., Roy, Kischkel, F., M., Dowd, P., Huang, A., Donahue, C., Sherwood, S., Baldwin, D., Godowski, P., Wood, W., Gurney, A., Hillan, K., Cohen, R., Goddard, A., Botstein, D., and Ashkenazi, A. Genomic amplification of a decoy receptor for Fas ligand in lung and colon cancer. *Nature* **396**, 699-703 (1998).

48. Mori, S., Marakami-Mori, K., Nakamura, S., Ashkenazi, A., and Bonavida, B. Sensitization of AIDS Kaposi's sarcoma cells to Apo-2 ligand-induced apoptosis by actinomycin D. *J. Immunol.* **162**, 5616-5623 (1999).

49. Gurney, A. Marsters, S., Huang, A., Pitti, R., Mark, M., Baldwin, D., Gray, A., Dowd, P., Brush, J., Heldens, S., Schow, P., Goddard, A., Wood, W., Baker, K., Godowski, P., and Ashkenazi, A. Identification of a new member of the tumor necrosis factor family and its receptor, a human ortholog of mouse GITR. *Curr. Biol.* **9**, 215-218 (1999).

50. Ashkenazi, A., Pai, R., Fong, s., Leung, S., Lawrence, D., Marsters, S., Blackie, C., Chang, L., McMurtrey, A., Hebert, A., DeForge, L., Khoumenis, I., Lewis, D., Harris, L., Bussiere, J., Koeppen, H., Shahrokh, Z., and Schwall, R. Safety and anti-tumor activity of recombinant soluble Apo2 ligand. *J. Clin. Invest.* **104**, 155-162 (1999).

51. Chuntharapai, A., Gibbs, V., Lu, J., Ow, A., Marsters, S., Ashkenazi, A., De Vos, A., Kim, K.J. Determination of residues involved in ligand binding and signal transmissiion in the human IFN-α receptor 2. *J. Immunol.* **163**, 766-773 (1999).

52. Johnsen, A.-C., Haux, J., Steinkjer, B., Nonstad, U., Egeberg, K., Sundan, A., Ashkenazi, A., and Espevik, T. Regulation of Apo2L/TRAIL expression in NK cells – involvement in NK cell-mediated cytotoxicity. *Cytokine* **11**, 664-672 (1999).

53. Roth, W., Isenmann, S., Naumann, U., Kugler, S., Bahr, M., Dichgans, J., Ashkenazi, A., and Weller, M. Eradication of intracranial human malignant glioma xenografts by Apo2L/TRAIL. *Biochem. Biophys. Res. Commun.* **265**, 479-483 (1999).

54. Hymowitz, S.G., Christinger, H.W., Fuh, G., Ultsch, M., O'Connell, M., Kelley, R.F., Ashkenazi, A. and de Vos, A.M. Triggering Cell Death: The Crystal Structure of Apo2L/TRAIL in a Complex with Death Receptor 5. *Molec. Cell* **4**, 563–571 (1999).

55. Hymowitz, S.G., O'Connel, M.P., Utsch, M.H., Hurst, A., Totpal, K., Ashkenazi, A., de Vos, A.M., Kelley, R.F. A unique zinc-binding site revealed by a high-resolution X-ray structure of homotrimeric Apo2L/TRAIL. *Biochemistry* **39**, 633-640 (2000).

56. Zhou, Q., Fukushima, P., DeGraff, W., Mitchell, J.B., Stetler-Stevenson, M., Ashkenazi, A., and Steeg, P.S. Radiation and the Apo2L/TRAIL apoptotic pathway preferentially inhibit the colonization of premalignant human breast cancer cells overexpressing cyclin D1. *Cancer Res.* **60**, 2611-2615 (2000).

57. Kischkel, F.C., Lawrence, D. A., Chuntharapai, A., Schow, P., Kim, J., and Ashkenazi, A. Apo2L/TRAIL-dependent recruitment of endogenous FADD and Caspase-8 to death receptors 4 and 5. *Immunity* **12**, 611-620 (2000).

58. Yan, M., Marsters, S.A., Grewal, I.S., Wang, H., *Ashkenazi, A., and *Dixit, V.M. Identification of a receptor for BlyS demonstrates a crucial role in humoral immunity. *Nature Immunol.* **1**, 37-41 (2000).

59. Marsters, S.A., Yan, M., Pitti, R.M., Haas, P.E., Dixit, V.M., and Ashkenazi, A. Interaction of the TNF homologues BLyS and APRIL with the TNF receptor homologues BCMA and TACI. *Curr. Biol.* **10**, 785-788 (2000).

60. Kischkel, F.C., and Ashkenazi, A. Combining enhanced metabolic labeling with immunoblotting to detect interactions of endogenous cellular proteins. *Biotechniques* **29**, 506-512 (2000).

61. Lawrence, D., Shahrokh, Z., Marsters, S., Achilles, K., Shih, D. Mounho, B., Hillan, K., Totpal, K. DeForge, L., Schow, P., Hooley, J., Sherwood, S., Pai, R., Leung, S., Khan, L., Gliniak, B., Bussiere, J., Smith, C., Strom, S., Kelley, S., Fox, J., Thomas, D., and Ashkenazi, A. Differential hepatocyte toxicity of recombinant Apo2L/TRAIL versions. *Nature Med.* **7**, 383-385 (2001).

62. Chuntharapai, A., Dodge, K., Grimmer, K., Schroeder, K., Martsters, S.A., Koeppen, H., Ashkenazi, A., and Kim, K.J. Isotype-dependent inhibition of tumor growth in vivo by monoclonal antibodies to death receptor 4. *J. Immunol.* **166**, 4891-4898 (2001).

63. Pollack, I.F., Erff, M., and Ashkenazi, A. Direct stimulation of apoptotic signaling by soluble Apo2L/tumor necrosis factor-related apoptosis-inducing ligand leads to selective killing of glioma cells. *Clin. Cancer Res.* **7**, 1362-1369 (2001).

64. Wang, H., Marsters, S.A., Baker, T., Chan, B., Lee, W.P., Fu, L., Tumas, D., Yan, M., Dixit, V.M., *Ashkenazi, A., and *Grewal, I.S. TACI-ligand interactions are required for T cell activation and collagen-induced arthritis in mice. *Nature Immunol.* 2, 632-637 (2001).

65. Kischkel, F.C., Lawrence, D. A., Tinel, A., Virmani, A., Schow, P., Gazdar, A., Blenis, J., Arnott, D., and Ashkenazi, A. Death receptor recruitment of endogenous caspase-10 and apoptosis initiation in the absence of caspase-8. *J. Biol. Chem.* 276, 46639-46646 (2001).

66. LeBlanc, H., Lawrence, D.A., Varfolomeev, E., Totpal, K., Morlan, J., Schow, P., Fong, S., Schwall, R., Sinicropi, D., and Ashkenazi, A Tumor cell resistance to death receptor induced apoptosis through mutational inactivation of the proapoptotitc Bcl-2 homolog Bax. *Nature Med.* 8, 274-281 (2002).

67. Miller, K., Meng, G., Liu, J., Hurst, A., Hsei, V., Wong, W-L., Ekert, R., Lawrence, D., Sherwood, S., DeForge, L., Gaudreault., Keller, G., Sliwkowski, M., Ashkenazi, A., and Presta, L. Design, Construction, and analyses of multivalent antibodies. *J. Immunol.* 170, 4854-4861 (2003).

68. Varfolomeev, E., Kischkel, F., Martin, F., Wanh, H., Lawrence, D., Olsson, C., Tom, L., Erickson, S., French, D., Schow, P., Grewal, I. and Ashkenazi, A. Immune system development in APRIL knockout mice. Submitted.

Review articles:

1. Ashkenazi, A., Peralta, E., Winslow, J., Ramachandran, J., and Capon, D., J. Functional role of muscarinic acetylcholine receptor subtype diversity. *Cold Spring Harbor Symposium on Quantitative Biology*. **LIII**, 263-272 (1988).
2. Ashkenazi, A., Peralta, E., Winslow, J., Ramachandran, J., and Capon, D. Functional diversity of muscarinic receptor subtypes in cellular signal transduction and growth. *Trends Pharmacol. Sci.* Dec Supplement, 12-21 (1989).
3. Chamow, S., Duliege, A., Ammann, A., Kahn, J., Allen, D., Eichberg, J., Byrn, R., Capon, D., Ward, R., and Ashkenazi, A. CD4 immunoadhesins in anti-HIV therapy: new developments. *Int. J. Cancer* Supplement 7, 69-72 (1992).
4. Ashkenazi, A., Capon, and D. Ward, R. Immunoadhesins. *Int. Rev. Immunol.* **10**, 217-225 (1993).
5. Ashkenazi, A., and Peralta, E. Muscarinic Receptors. In *Handbook of Receptors and Channels*. (S. Peroutka, ed.), CRC Press, Boca Raton, Vol. **I**, p. 1-27, (1994).
6. Krantz, S. B., Means, R. T., Jr., Lina, J., Marsters, S. A., and Ashkenazi, A. Inhibition of erythroid colony formation in vitro by gamma interferon. In *Molecular Biology of Hematopoiesis* (N. Abraham, R. Shadduck, A. Levine F. Takaku, eds.) Intercept Ltd. Paris, Vol. **3**, p. 135-147 (1994).
7. Ashkenazi, A. Cytokine neutralization as a potential therapeutic approach for SIRS and shock. *J. Biotechnology in Healthcare* **1**, 197-206 (1994).
8. Ashkenazi, A., and Chamow, S. M. Immunoadhesins: an alternative to human monoclonal antibodies. *Immunomethods: A companion to Methods in Enzimology* **8**, 104-115 (1995).
9. Chamow, S., and Ashkenazi, A. Immunoadhesins: Principles and Applications. *Trends Biotech.* **14**, 52-60 (1996).
10. Ashkenazi, A., and Chamow, S. M. Immunoadhesins as research tools and therapeutic agents. *Curr. Opin. Immunol.* **9**, 195-200 (1997).
11. Ashkenazi, A., and Dixit, V. Death receptors: signaling and modulation. *Science* **281**, 1305-1308 (1998).
12. Ashkenazi, A., and Dixit, V. Apoptosis control by death and decoy receptors. *Curr. Opin. Cell. Biol.* **11**, 255-260 (1999).

13. Ashkenazi, A. Chapters on Apo2L/TRAIL; DR4, DR5, DcR1, DcR2; and DcR3. Online Cytokine Handbook (www.apnet.com/cytokinereference/).
14. Ashkenazi, A. Targeting death and decoy receptors of the tumor necrosis factor superfamily. *Nature Rev. Cancer* **2**, 420-430 (2002).
15. LeBlanc, H. and Ashkenazi, A. Apoptosis signaling by Apo2L/TRAIL. *Cell Death and Differentiation* 10, 66-75 (2003).
16. Almasan, A. and Ashkenazi, A. Apo2L/TRAIL: apoptosis signaling, biology, and potential for cancer therapy. Cytokine and Growth Factor Reviews 14, 337-348 (2003).

Book:

Antibody Fusion Proteins (Chamow, S., and Ashkenazi, A., eds., John Wiley and Sons Inc.) (1999).

Talks:

1. Resistance of primary HIV isolates to CD4 is independent of CD4-gp120 binding affinity. UCSD Symposium, HIV Disease: Pathogenesis and Therapy. Greenelefe, FL, March 1991.
2. Use of immuno-hybrids to extend the half-life of receptors. IBC conference on Biopharmaceutical Halflife Extension. New Orleans, LA, June 1992.
3. Results with TNF receptor Immunoadhesins for the Treatment of Sepsis. IBC conference on Endotoxemia and Sepsis. Philadelphia, PA, June 1992.
4. Immunoadhesins: an alternative to human antibodies. IBC conference on Antibody Engineering. San Diego, CA, December 1993.
5. Tumor necrosis factor receptor: a potential therapeutic for human septic shock. American Society for Microbiology Meeting, Atlanta, GA, May 1993.
6. Protective efficiacy of TNF receptor immunoadhesin vs anti-TNF monoclonal antibody in a rat model for endotoxic shock. 5th International Congress on TNF. Asilomar, CA, May 1994.
7. Interferon-γ signals via a multisubunit receptor complex that contains two types of polypeptide chain. American Association of Immunologists Conference. San Franciso, CA, July 1995.
8. Immunoadhesins: Principles and Applications. Gordon Research Conference on Drug Delivery in Biology and Medicine. Ventura, CA, February 1996.

9. Apo-2 Ligand, a new member of the TNF family that induces apoptosis in tumor cells. Cambridge Symposium on TNF and Related Cytokines in Treatment of Cancer. Hilton-Head, NC, March 1996.
10. Induction of apoptosis by Apo2 Ligand. American Society for Biochemistry and Molecular Biology, Symposium on Growth Factors and Cytokine Receptors. New Orleans, LA, June, 1996.
11. Apo2 ligand, an extracellular trigger of apoptosis. 2nd Clontech Symposium, Palo Alto, CA, October 1996.
12. Regulation of apoptosis by members of the TNF ligand and receptor families. Stanford University School of Medicine, Palo Alto, CA, December 1996.
13. Apo-3: anovel receptor that regulates cell death and inflammation. 4th International Congress on Immune Consequences of Trauma, Shock, and Sepsis. Munich, Germany, March 1997.
14. New members of the TNF ligand and receptor families that regulate apoptosis, inflammation, and immunity. UCLA School of Medicine, LA, CA, March 1997.
15. Immunoadhesins: an alternative to monoclonal antibodies. 5th World Conference on Bispecific Antibodies. Volendam, Holland, June 1997.
16. Control of Apo2L signaling. Cold Spring Harbor Laboratory Symposium on Programmed Cell Death. Cold Spring Harbor, New York. September, 1997.
17. Chairman and speaker, Apoptosis Signaling session. IBC's 4th Annual Conference on Apoptosis. San Diego, CA., October 1997.
18. Control of Apo2L signaling by death and decoy receptors. American Association for the Advancement of Science. Philladelphia, PA, February 1998.
19. Apo2 ligand and its receptors. American Society of Immunologists. San Francisco, CA, April 1998.
20. Death receptors and ligands. 7th International TNF Congress. Cape Cod, MA, May 1998.
21. Apo2L as a potential therapeutic for cancer. UCLA School of Medicine. LA, CA, June 1998.
22. Apo2L as a potential therapeutic for cancer. Gordon Research Conference on Cancer Chemotherapy. New London, NH, July 1998.
23. Control of apoptosis by Apo2L. Endocrine Society Conference, Stevenson, WA, August 1998.
24. Control of apoptosis by Apo2L. International Cytokine Society Conference, Jerusalem, Israel, October 1998.

25. Apoptosis control by death and decoy receptors. American Association for Cancer Research Conference, Whistler, BC, Canada, March 1999.
26. Apoptosis control by death and decoy receptors. American Society for Biochemistry and Molecular Biology Conference, San Francisco, CA, May 1999.
27. Apoptosis control by death and decoy receptors. Gordon Research Conference on Apoptosis, New London, NH, June 1999.
28. Apoptosis control by death and decoy receptors. Arthritis Foundation Research Conference, Alexandria GA, Aug 1999.
29. Safety and anti-tumor activity of recombinant soluble Apo2L/TRAIL. Cold Spring Harbor Laboratory Symposium on Programmed Cell Death. . Cold Spring Harbor, NY, September 1999.
30. The Apo2L/TRAIL system: therapeutic potential. American Association for Cancer Research, Lake Tahoe, NV, Feb 2000.
31. Apoptosis and cancer therapy. Stanford University School of Medicine, Stanford, CA, Mar 2000.
32. Apoptosis and cancer therapy. University of Pennsylvania School of Medicine, Philladelphia, PA, Apr 2000.
33. Apoptosis signaling by Apo2L/TRAIL. International Congress on TNF. Trondheim, Norway, May 2000.
34. The Apo2L/TRAIL system: therapeutic potential. Cap-CURE summit meeting. Santa Monica, CA, June 2000.
35. The Apo2L/TRAIL system: therapeutic potential. MD Anderson Cancer Center. Houston, TX, June 2000.
36. Apoptosis signaling by Apo2L/TRAIL. The Protein Society, 14th Symposium. San Diego, CA, August 2000.
37. Anti-tumor activity of Apo2L/TRAIL. AAPS annual meeting. Indianapolis, IN Aug 2000.
38. Apoptosis signaling and anti-cancer potential of Apo2L/TRAIL. Cancer Research Institute, UC San Francisco, CA, September 2000.
39. Apoptosis signaling by Apo2L/TRAIL. Kenote address, TNF family Minisymposium, NIH. Bethesda, MD, September 2000.
40. Death receptors: signaling and modulation. Keystone symposium on the Molecular basis of cancer. Taos, NM, Jan 2001.
41. Preclinical studies of Apo2L/TRAIL in cancer. Symposium on Targeted therapies in the treatment of lung cancer. Aspen, CO, Jan 2001.

42. Apoptosis signaling by Apo2L/TRAIL. Wiezmann Institute of Science, Rehovot, Israel, March 2001.
43. Apo2L/TRAIL: Apoptosis signaling and potential for cancer therapy. Weizmann Institute of Science, Rehovot, Israel, March 2001.
44. Targeting death receptors in cancer with Apo2L/TRAIL. Cell Death and Disease conference, North Falmouth, MA, Jun 2001.
45. Targeting death receptors in cancer with Apo2L/TRAIL. Biotechnology Organization conference, San Diego, CA, Jun 2001.
46. Apo2L/TRAIL signaling and apoptosis resistance mechanisms. Gordon Research Conference on Apoptosis, Oxford, UK, July 2001.
47. Apo2L/TRAIL signaling and apoptosis resistance mechanisms. Cleveland Clinic Foundation, Cleveland, OH, Oct 2001.
48. Apoptosis signaling by death receptors: overview. International Society for Interferon and Cytokine Research conference, Cleveland, OH, Oct 2001.
49. Apoptosis signaling by death receptors. American Society of Nephrology Conference. San Francisco, CA, Oct 2001.
50. Targeting death receptors in cancer. Apoptosis: commercial opportunities. San Diego, CA, Apr 2002.
51. Apo2L/TRAIL signaling and apoptosis resistance mechanisms. Kimmel Cancer Research Center, Johns Hopkins University, Baltimore MD. May 2002.
52. Apoptosis control by Apo2L/TRAIL. (Keynote Address) University of Alabama Cancer Center Retreat, Birmingham, Ab. October 2002.
53. Apoptosis signaling by Apo2L/TRAIL. (Session co-chair) TNF international conference. San Diego, CA. October 2002.
54. Apoptosis signaling by Apo2L/TRAIL. Swiss Institute for Cancer Research (ISREC). Lausanne, Swizerland. Jan 2003.
55. Apoptosis induction with Apo2L/TRAIL. Conference on New Targets and Innovative Strategies in Cancer Treatment. Monte Carlo. February 2003.
56. Apoptosis signaling by Apo2L/TRAIL. Hermelin Brain Tumor Center Symposium on Apoptosis. Detroit, MI. April 2003.
57. Targeting apoptosis through death receptors. Sixth Annual Conference on Targeted Therapies in the Treatment of Breast Cancer. Kona, Hawaii. July 2003.
58. Targeting apoptosis through death receptors. Second International Conference on Targeted Cancer Therapy. Washington, DC. Aug 2003.

Issued Patents:

1. Ashkenazi, A., Chamow, S. and Kogan, T. Carbohydrate-directed crosslinking reagents. US patent 5,329,028 (Jul 12, 1994).
2. Ashkenazi, A., Chamow, S. and Kogan, T. Carbohydrate-directed crosslinking reagents. US patent 5,605,791 (Feb 25, 1997).
3. Ashkenazi, A., Chamow, S. and Kogan, T. Carbohydrate-directed crosslinking reagents. US patent 5,889,155 (Jul 27, 1999).
4. Ashkenazi, A., APO-2 Ligand. US patent 6,030,945 (Feb 29, 2000).
5. Ashkenazi, A., Chuntharapai, A., Kim, J., APO-2 ligand antibodies. US patent 6, 046, 048 (Apr 4, 2000).
6. Ashkenazi, A., Chamow, S. and Kogan, T. Carbohydrate-directed crosslinking reagents. US patent 6,124,435 (Sep 26, 2000).
7. Ashkenazi, A., Chuntharapai, A., Kim, J., Method for making monoclonal and cross-reactive antibodies. US patent 6,252,050 (Jun 26, 2001).
8. Ashkenazi, A. APO-2 Receptor. US patent 6,342,369 (Jan 29, 2002).
9. Ashkenazi, A. Fong, S., Goddard, A., Gurney, A., Napier, M., Tumas, D., Wood, W. A-33 polypeptides. US patent 6,410,708 (Jun 25, 2002).
10. Ashkenazi, A. APO-3 Receptor. US patent 6,462,176 B1 (Oct 8, 2002).
11. Ashkenazi, A. APO-2LI and APO-3 polypeptide antibodies. US patent 6,469,144 B1 (Oct 22, 2002).
12. Ashkenazi, A., Chamow, S. and Kogan, T. Carbohydrate-directed crosslinking reagents. US patent 6,582,928B1 (Jun 24, 2003).

TECHNICAL UPDATE

FROM YOUR LABORATORY SERVICES PROVIDER

HER-2/neu Breast Cancer Predictive Testing

Julie Sanford Hanna, Ph.D. and Dan Mornin, M.D.

EACH YEAR, OVER 182,000 WOMEN in the United States are diagnosed with breast cancer, and approximately 45,000 die of the disease.[1] Incidence appears to be increasing in the United States at a rate of roughly 2% per year. The reasons for the increase are unclear, but non-genetic risk factors appear to play a large role.[2]

Five-year survival rates range from approximately 65%-85%, depending on demographic group, with a significant percentage of women experiencing recurrence of their cancer within 10 years of diagnosis. One of the factors most predictive for recurrence once a diagnosis of breast cancer has been made is the number of axillary lymph nodes to which tumor has metastasized. Most node-positive women are given adjuvant therapy, which increases their survival. However, 20%-30% of patients without axillary node involvement also develop recurrent disease, and the difficulty lies in how to identify this high-risk subset of patients. These patients could benefit from increased surveillance, early intervention, and treatment.

Prognostic markers currently used in breast cancer recurrence prediction include tumor size, histological grade, steroid hormone receptor status, DNA ploidy, proliferative index, and cathepsin D status. Expression of growth factor receptors and over-expression of the HER-2/neu oncogene have also been identified as having value regarding treatment regimen and prognosis.

HER-2/neu (also known as c-erbB2) is an oncogene that encodes a transmembrane glycoprotein that is homologous to, but distinct from, the epidermal growth factor receptor. Numerous studies have indicated that high levels of expression of this protein are associated with rapid tumor growth, certain forms of therapy resistance, and shorter disease-free survival. The gene has been shown to be amplified and/or overexpressed in 10%-30% of invasive breast cancers and in 40%-60% of intraductal breast carcinoma.[3]

There are two distinct FDA-approved methods by which HER-2/neu status can be evaluated: immunohistochemistry (IHC, HercepTest™) and FISH (fluorescent in situ hybridization, PathVysion™ Kit). Both methods can be performed on archived and current specimens. The first method allows visual assessment of the amount of HER-2/neu protein present on the cell membrane. The latter method allows direct quantification of the level of gene amplification present in the tumor, enabling differentiation between low- versus high-amplification. At least one study has demonstrated a difference in recurrence risk in women younger than 40 years of age for low- versus high-amplified tumors (54.5% compared to 85.7%); this is compared to a recurrence rate of 16.7% for patients with no HER-2/neu gene amplification.[4] HER-2/neu status may be particularly important to establish in women with small ($\leq$1 cm) tumor size.

The choice of methodology for determination of HER-2/neu status depends in part on the clinical setting. FDA approval for the Vysis FISH test was granted based on clinical trials involving 1549 node-positive patients. Patients received one of three different treatments consisting of different doses of cyclophosphamide, Adriamycin, and 5-fluorouracil (CAF). The study showed that patients with amplified HER-2/neu benefited from treatment with higher doses of adriamycin-based therapy, while those with normal HER-2/neu levels did not. The study therefore identified a sub-set of women, who because they did not benefit from more aggressive treatment, did not need to be exposed to the associated side effects. In addition, other evidence indicates that HER-2/neu amplification in node-negative patients can be used as an independent prognostic indicator for early recurrence, recurrent disease at any time and disease-related death.[5] Demonstration of HER-2/neu gene amplification by FISH has also been shown to be of value in predicting response to chemotherapy in stage-2 breast cancer patients.

Selection of patients for Herceptin© (Trastuzumab) monoclonal antibody therapy, however, is based upon demonstration of HER-2/neu protein overexpression using HercepTest™. Studies using Herceptin© in patients with metastatic breast cancer show an increase in time to disease progression, increased response rate to chemotherapeutic agents and a small increase in overall survival rate. The FISH assays have not yet been approved for this purpose, and studies looking at response to Herceptin© in patients with or without gene amplification status determined by FISH are in progress.

In general, FISH and IHC results correlate well. However, subsets of tumors are found which show discordant results; i.e., protein overexpression without gene amplification or lack of protein overexpression with gene amplification. The clinical significance of such results is unclear. Based on the above considerations, HER-2/neu testing at SHMC/PAML will utilize immunohistochemistry (HercepTest©) as a screen, followed by FISH in IHC-negative cases. Alternatively, either method may be ordered individually depending on the clinical setting or clinician preference.

Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors

Jonathan R. Pollack*†‡, Therese Sørlie§, Charles M. Perou¶, Christian A. Rees‖**, Stefanie S. Jeffrey††, Per E. Lonning‡‡, Robert Tibshirani§§, David Botstein‖, Anne-Lise Børresen-Dale§, and Patrick O. Brown†¶¶

Departments of *Pathology, ‖Genetics, ††Surgery, §§Health Research and Policy, and ¶¶Biochemistry, and †Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305; §Department of Genetics, Norwegian Radium Hospital, Montebello, N-0310 Oslo, Norway; ‡‡Department of Medicine (Oncology), Haukeland University Hospital, N-5021 Bergen, Norway; and ¶Department of Genetics and Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599

Contributed by Patrick O. Brown, August 6, 2002

Genomic DNA copy number alterations are key genetic events in the development and progression of human cancers. Here we report a genome-wide microarray comparative genomic hybridization (array CGH) analysis of DNA copy number variation in a series of primary human breast tumors. We have profiled DNA copy number alteration across 6,691 mapped human genes, in 44 predominantly advanced, primary breast tumors and 10 breast cancer cell lines. While the overall patterns of DNA amplification and deletion corroborate previous cytogenetic studies, the high-resolution (gene-by-gene) mapping of amplicon boundaries and the quantitative analysis of amplicon shape provide significant improvement in the localization of candidate oncogenes. Parallel microarray measurements of mRNA levels reveal the remarkable degree to which variation in gene copy number contributes to variation in gene expression in tumor cells. Specifically, we find that 62% of highly amplified genes show moderately or highly elevated expression, that DNA copy number influences gene expression across a wide range of DNA copy number alterations (deletion, low-, mid- and high-level amplification), that on average, a 2-fold change in DNA copy number is associated with a corresponding 1.5-fold change in mRNA levels, and that overall, at least 12% of all the variation in gene expression among the breast tumors is directly attributable to underlying variation in gene copy number. These findings provide evidence that widespread DNA copy number alteration can lead directly to global deregulation of gene expression, which may contribute to the development or progression of cancer.

Conventional cytogenetic techniques, including comparative genomic hybridization (CGH) (1), have led to the identification of a number of recurrent regions of DNA copy number alteration in breast cancer cell lines and tumors (2–4). While some of these regions contain known or candidate oncogenes [e.g., FGFR1 (8p11), MYC (8q24), CCND1 (11q13), ERBB2 (17q12), and ZNF217 (20q13)] and tumor suppressor genes [RB1 (13q14) and TP53 (17p13)], the relevant gene(s) within other regions (e.g., gain of 1q, 8q22, and 17q22–24, and loss of 8p) remain to be identified. A high-resolution genome-wide map, delineating the boundaries of DNA copy number alterations in tumors, should facilitate the localization and identification of oncogenes and tumor suppressor genes in breast cancer. In this study, we have created such a map, using array-based CGH (5–7) to profile DNA copy number alteration in a series of breast cancer cell lines and primary tumors.

An unresolved question is the extent to which the widespread DNA copy number changes that we and others have identified in breast tumors alter expression of genes within involved regions. Because we had measured mRNA levels in parallel in the same samples (8), using the same DNA microarrays, we had an opportunity to explore on a genomic scale the relationship between DNA copy number changes and gene expression. From this analysis, we have identified a significant impact of widespread DNA copy number alteration on the transcriptional programs of breast tumors.

Materials and Methods

Tumors and Cell Lines. Primary breast tumors were predominantly large (>3 cm), intermediate-grade, infiltrating ductal carcinomas, with more than 50% being lymph node positive. The fraction of tumor cells within specimens averaged at least 50%. Details of individual tumors have been published (8, 9), and are summarized in Table 1, which is published as supporting information on the PNAS web site, www.pnas.org. Breast cancer cell lines were obtained from the American Type Culture Collection. Genomic DNA was isolated either using Qiagen genomic DNA columns, or by phenol/chloroform extraction followed by ethanol precipitation.

DNA Labeling and Microarray Hybridizations. Genomic DNA labeling and hybridizations were performed essentially as described in Pollack *et al.* (7), with slight modifications. Two micrograms of DNA was labeled in a total volume of 50 microliters and the volumes of all reagents were adjusted accordingly. "Test" DNA (from tumors and cell lines) was fluorescently labeled (Cy5) and hybridized to a human cDNA microarray containing 6,691 different mapped human genes (i.e., UniGene clusters). The "reference" (labeled with Cy3) for each hybridization was normal female leukocyte DNA from a single donor. The fabrication of cDNA microarrays and the labeling and hybridization of mRNA samples have been described (8).

Data Analysis and Map Positions. Hybridized arrays were scanned on a GenePix scanner (Axon Instruments, Foster City, CA), and fluorescence ratios (test/reference) calculated using SCANALYZE software (available at http://rana.lbl.gov). Fluorescence ratios were normalized for each array by setting the average log fluorescence ratio for all array elements equal to 0. Measurements with fluorescence intensities more than 20% above background were considered reliable. DNA copy number profiles that deviated significantly from background ratios measured in normal genomic DNA control hybridizations were interpreted as evidence of real DNA copy number alteration (see *Estimating Significance of Altered Fluorescence Ratios* in the supporting information). When indicated, DNA copy number profiles are displayed as a moving average (symmetric 5-nearest neighbors). Map positions for arrayed human cDNAs were assigned by

Abbreviation: CGH, comparative genomic hybridization.

†To whom reprint requests should be addressed at: Department of Pathology, Stanford University School of Medicine, CCSR Building, Room 3245A, 269 Campus Drive, Stanford, CA 94305-5176. E-mail: pollack1@stanford.edu.

**Present address: Zyomyx Inc., Hayward, CA 94545.

GENETICS

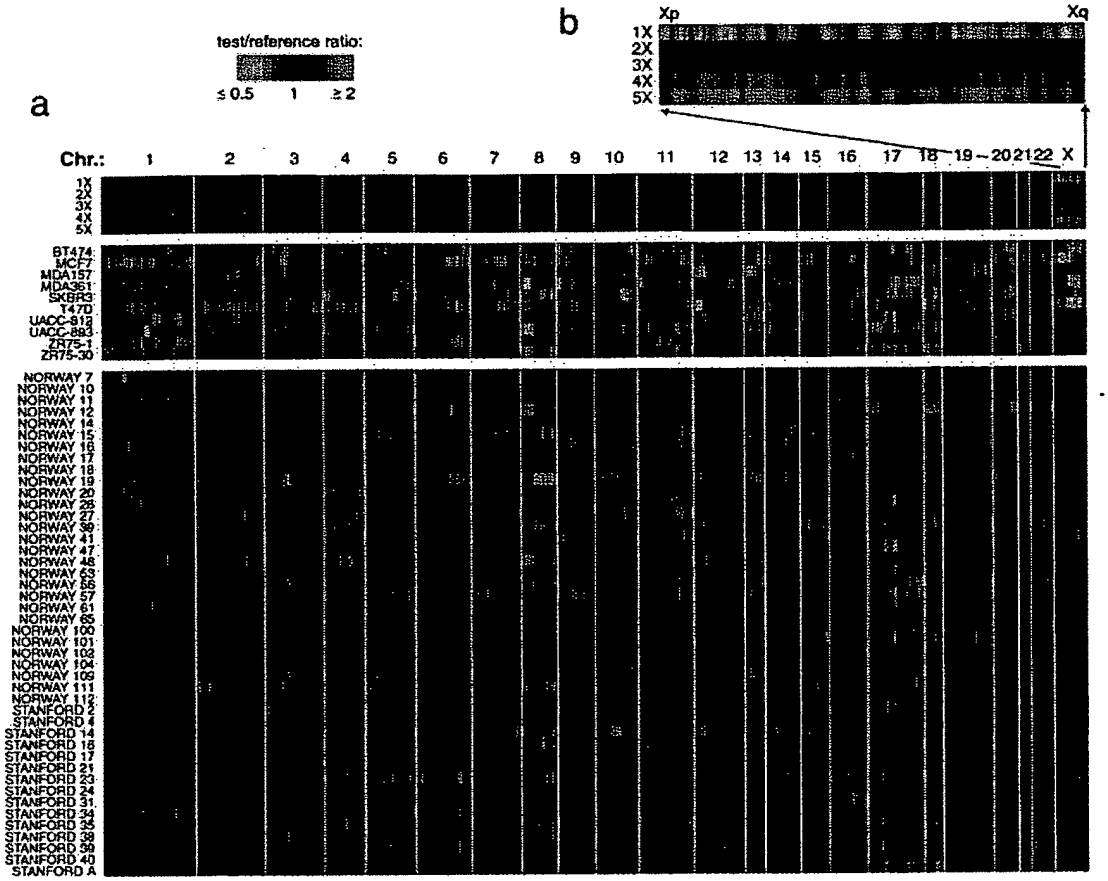


Fig. 1. Genome-wide measurement of DNA copy number alteration by array CGH. (*a*) DNA copy number profiles are illustrated for cell lines containing different numbers of X chromosomes, for breast cancer cell lines, and for breast tumors. Each row represents a different cell line or tumor, and each column represents one of 6,691 different mapped human genes present on the microarray, ordered by genome map position from 1pter through Xqter. Moving average (symmetric 5-nearest neighbors) fluorescence ratios (test/reference) are depicted using a $\log_2$-based pseudocolor scale (indicated), such that red luminescence reflects fold-amplification, green luminescence reflects fold-deletion, and black indicates no change (gray indicates poorly measured data). (*b*) Enlarged view of DNA copy number profiles across the X chromosome, shown for cell lines containing different numbers of X chromosomes.

identifying the starting position of the best and longest match of any DNA sequence represented in the corresponding UniGene cluster (10) against the "Golden Path" genome assembly (http://genome.ucsc.edu/; Oct 7, 2000 Freeze). For UniGene clusters represented by multiple arrayed elements, mean fluorescence ratios (for all elements representing the same UniGene cluster) are reported. For mRNA measurements, fluorescence ratios are "mean-centered" (i.e., reported relative to the mean ratio across the 44 tumor samples). The data set described here can be accessed in its entirety in the supporting information.

Results

We performed CGH on 44 predominantly locally advanced, primary breast tumors and 10 breast cancer cell lines, using cDNA microarrays containing 6,691 different mapped human genes (Fig. 1*a*; also see *Materials and Methods* for details of microarray hybridizations). To take full advantage of the improved spatial resolution of array CGH, we ordered (fluorescence ratios for) the 6,691 cDNAs according to the "Golden Path" (http://genome.ucsc.edu/) genome assembly of the draft human genome sequences (11). In so doing, arrayed cDNAs not only themselves represent genes of potential interest (e.g., candidate oncogenes within amplicons), but also provide precise genetic landmarks for chromosomal regions of amplification and deletion. Parallel analysis of DNA from cell lines containing different numbers of X chromosomes (Fig. 1*b*), as we did before (7), demonstrated the sensitivity of our method to detect single-copy loss (45, XO), and 1.5- (47,XXX), 2- (48,XXXX), or 2.5-fold (49,XXXXX) gains (also see Fig. 5, which is published as supporting information on the PNAS web site). Fluorescence ratios were linearly proportional to copy number ratios, which were slightly underestimated, in agreement with previous observations (7). Numerous DNA copy number alterations were evident in both the breast cancer cell lines and primary tumors (Fig. 1*a*), detected in the tumors despite the presence of euploid non-tumor cell types; the magnitudes of the observed changes were generally lower in the tumor samples. DNA copy-number alterations were found in every cancer cell line and tumor, and on every human chromosome in at least one sample. Recurrent regions of DNA copy number gain and loss were readily identifiable. For example, gains within 1q, 8q, 17q, and 20q were observed in a high proportion of breast cancer cell lines/tumors (90%/69%, 100%/47%, 100%/60%, and 90%/44%, respectively), as were losses within 1p, 3p, 8p, and 13q (80%/24%, 80%/22%, 80%/22%, and 70%/18%, respectively), consistent with published cytogenetic studies (refs. 2–4; a complete listing of gains/losses is provided in Tables 2 and 3, which are published as supporting information on the PNAS web site). The total

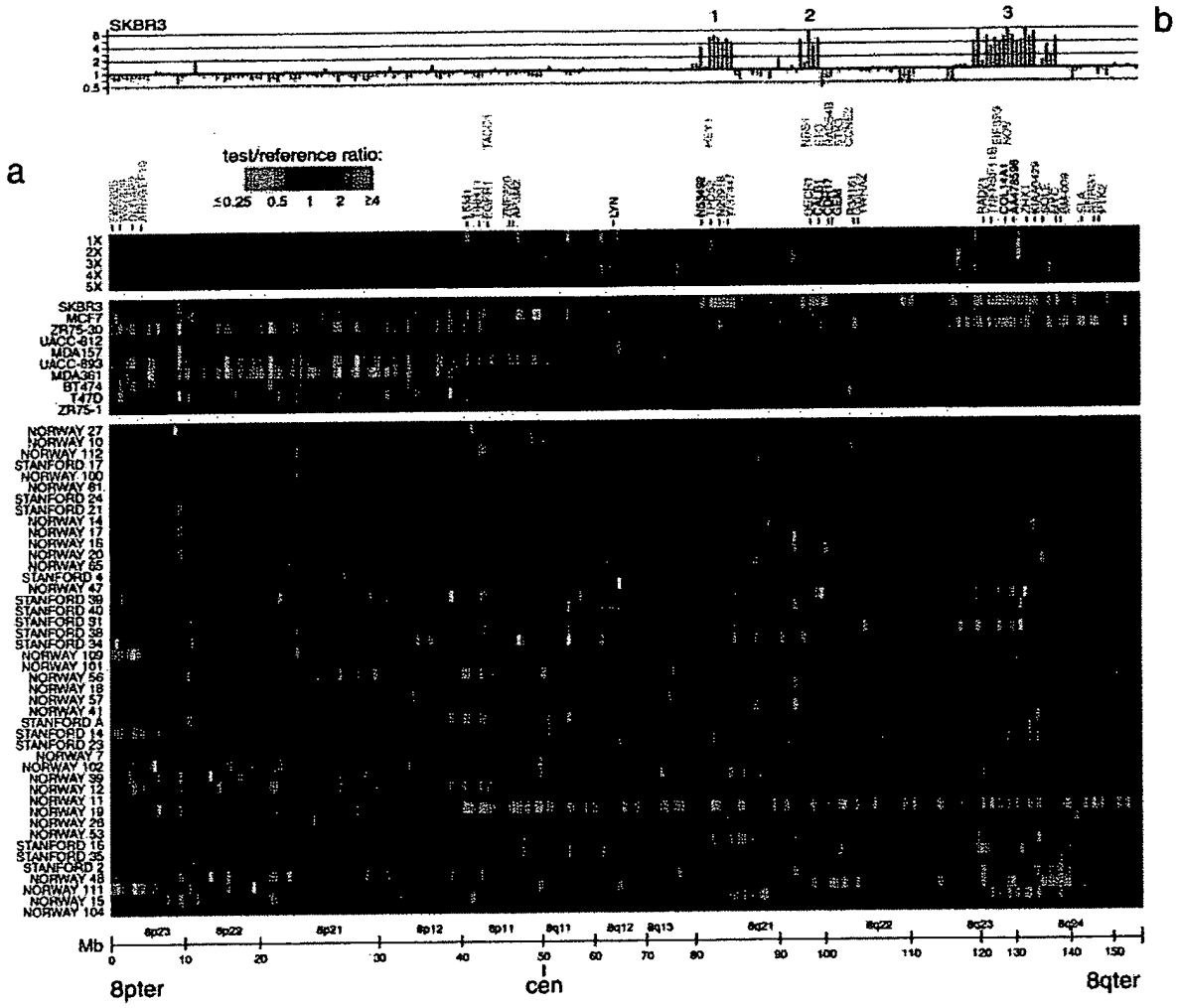


Fig. 2. DNA copy number alteration across chromosome 8 by array CGH. (*a*) DNA copy number profiles are illustrated for cell lines containing different numbers of X chromosomes, for breast cancer cell lines, and for breast tumors. Breast cancer cell lines and tumors are separately ordered by hierarchical clustering to highlight recurrent copy number changes. The 241 genes present on the microarrays and mapping to chromosome 8 are ordered by position along the chromosome. Fluorescence ratios (test/reference) are depicted by a $\log_2$ pseudocolor scale (indicated). Selected genes are indicated with color-coded text (red, increased; green, decreased; black, no change; gray, not well measured) to reflect correspondingly altered mRNA levels (observed in the majority of the subset of samples displaying the DNA copy number change). The map positions for genes of interest that are not represented on the microarray are indicated in the row above those genes represented on the array. (*b*) Graphical display of DNA copy number profile for breast cancer cell line SKBR3. Fluorescence ratios (tumor/normal) are plotted on a $\log_2$ scale for chromosome 8 genes, ordered along the chromosome.

number of genomic alterations (gains and losses) was found to be significantly higher in breast tumors that were high grade (P = 0.008), consistent with published CGH data (3), estrogen receptor negative (P = 0.04), and harboring TP53 mutations (P = 0.0006) (see Table 4, which is published as supporting information on the PNAS web site).

The improved spatial resolution of our array CGH analysis is illustrated for chromosome 8, which displayed extensive DNA copy number alteration in our series. A detailed view of the variation in the copy number of 241 genes mapping to chromosome 8 revealed multiple regions of recurrent amplification; each of these potentially harbors a different known or previously uncharacterized oncogene (Fig. 2*a*). The complexity of amplicon structure is most easily appreciated in the breast cancer cell line SKBR3. Although a conventional CGH analysis of 8q in SKBR3 identified only two distinct regions of amplification (12), we observed three distinct regions of high-level amplification (labeled 1–3 in Fig. 2*b*). For each of these regions we can define the boundaries of the interval recurrently amplified in the tumors we examined; in each case, known or plausible candidate oncogenes can be identified (a description of these regions, as well as the recurrently amplified regions on chromosomes 17 and 20, can be found in Figs. 6 and 7, which are published as supporting information on the PNAS web site).

For a subset of breast cancer cell lines and tumors (4 and 37, respectively), and a subset of arrayed genes (6,095), mRNA levels were quantitatively measured in parallel by using cDNA microarrays (8). The parallel assessment of mRNA levels is useful in the interpretation of DNA copy number changes. For example, the highly amplified genes that are also highly expressed are the strongest candidate oncogenes within an amplicon. Perhaps more significantly, our parallel analysis of DNA copy number changes and mRNA levels provides us the opportunity to assess the global impact of widespread DNA copy number alteration on gene expression in tumor cells.

A strong influence of DNA copy number on gene expression is evident in an examination of the pseudocolor representations

GENETICS

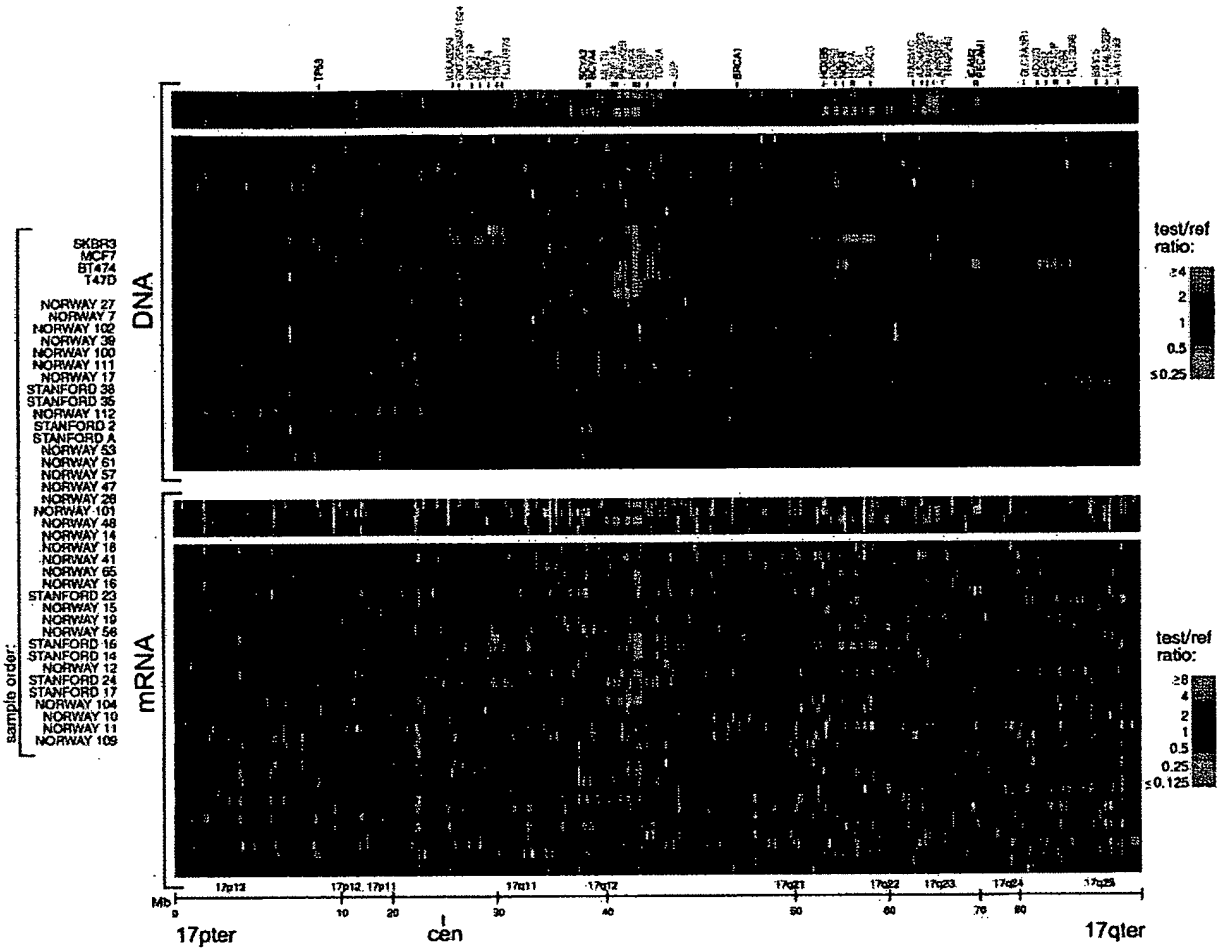


Fig. 3. Concordance between DNA copy number and gene expression across chromosome 17. DNA copy number alteration (*Upper*) and mRNA levels (*Lower*) are illustrated for breast cancer cell lines and tumors. Breast cancer cell lines and tumors are separately ordered by hierarchical clustering (*Upper*), and the identical sample order is maintained (*Lower*). The 354 genes present on the microarrays and mapping to chromosome 17, and for which both DNA copy number and mRNA levels were determined, are ordered by position along the chromosome; selected genes are indicated in color-coded text (see Fig. 2 legend). Fluorescence ratios (test/reference) are depicted by separate $\log_2$ pseudocolor scales (indicated).

of DNA copy number and mRNA levels for genes on chromosome 17 (Fig. 3). The overall patterns of gene amplification and elevated gene expression are quite concordant; i.e., a significant fraction of highly amplified genes appear to be correspondingly highly expressed. The concordance between high-level amplification and increased gene expression is not restricted to chromosome 17. Genome-wide, of 117 high-level DNA amplifications (fluorescence ratios >4, and representing 91 different genes), 62% (representing 54 different genes; see Table 5, which is published as supporting information on the PNAS web site) are found associated with at least moderately elevated mRNA levels (mean-centered fluorescence ratios >2), and 42% (representing 36 different genes) are found associated with comparably highly elevated mRNA levels (mean-centered fluorescence ratios >4).

To determine the extent to which DNA deletion and lower-level amplification (in addition to high-level amplification) are also associated with corresponding alterations in mRNA levels, we performed three separate analyses on the complete data set (4 cell lines and 37 tumors, across 6,095 genes). First, we determined the average mRNA levels for each of five classes of genes, representing DNA deletion, no change, and low-, medium-, and high-level amplification (Fig. 4*a*). For both the breast cancer cell lines and tumors, average mRNA levels tracked with DNA copy number across all five classes, in a statistically significant fashion (P values for pair-wise Student's t tests comparing adjacent classes: cell lines, 4×10^{-49}, 1×10^{-49}, 5×10^{-5}, 1×10^{-2}; tumors, 1×10^{-43}, 1×10^{-214}, 5×10^{-41}, 1×10^{-4}). A linear regression of the average log(DNA copy number), for each class, against average log(mRNA level) demonstrated that on average, a 2-fold change in DNA copy number was accompanied by 1.4- and 1.5-fold changes in mRNA level for the breast cancer cell lines and tumors, respectively (Fig. 4*a*, regression line not shown). Second, we characterized the distribution of the 6,095 correlations between DNA copy number and mRNA level, each across the 37 tumor samples (Fig. 4*b*). The distribution of correlations forms a normal-shaped curve, but with the peak markedly shifted in the positive direction from zero. This shift is statistically significant, as evidenced in a plot of observed vs. expected correlations (Fig. 4*c*), and reflects a pervasive global influence of DNA copy number alterations on gene expression. Notably, the highest correlations between DNA copy number and mRNA level (the right tail of the distribution in Fig. 4*b*) comprise both amplified and deleted genes (data not shown). Third, we used a linear regression model to estimate the fraction of all variation measured in mRNA levels among the 37

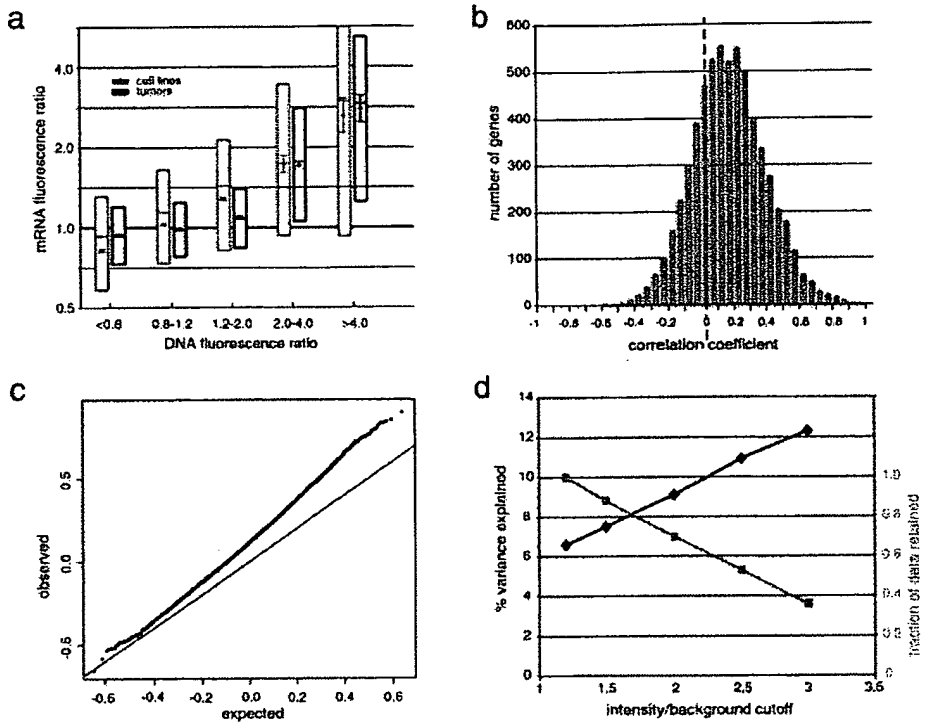


Fig. 4. Genome-wide influence of DNA copy number alterations on mRNA levels. (a) For breast cancer cell lines (gray) and tumor samples (black), both mean-centered mRNA fluorescence ratio (log_2 scale) quartiles (box plots indicate 25th, 50th, and 75th percentile) and averages (diamonds; Y-value error bars indicate standard errors of the mean) are plotted for each of five classes of genes, representing DNA deletion (tumor/normal ratio < 0.8), no change (0.8–1.2), low- (1.2–2), medium- (2–4), and high-level (>4) amplification. *P* values for pair-wise Student's *t* tests, comparing averages between adjacent classes (moving left to right), are 4×10^{-49}, 1×10^{-49}, 5×10^{-5}, 1×10^{-2} (cell lines), and 1×10^{-43}, 1×10^{-214}, 5×10^{-41}, 1×10^{-4} (tumors). (*b*) Distribution of correlations between DNA copy number and mRNA levels, for 6,095 different human genes across 37 breast tumor samples. (*c*) Plot of observed versus expected correlation coefficients. The expected values were obtained by randomization of the sample labels in the DNA copy number data set. The line of unity is indicated. (*d*) Percent variance in gene expression (among tumors) directly explained by variation in gene copy number. Percent variance explained (black line) and fraction of data retained (gray line) are plotted for different fluorescence intensity/background (a rough surrogate for signal/noise) cutoff values. Fraction of data retained is relative to the 1.2 intensity/background cutoff. Details of the linear regression model used to estimate the fraction of variation in gene expression attributable to underlying DNA copy number alteration can be found in the supporting information (see *Estimating the Fraction of Variation in Gene Expression Attributable to Underlying DNA Copy Number Alteration*).

tumors that could be attributed to underlying variation in DNA copy number. From this analysis, we estimate that, overall, about 7% of all of the observed variation in mRNA levels can be explained directly by variation in copy number of the altered genes (Fig. 4*d*). We can reduce the effects of experimental measurement error on this estimate by using only that fraction of the data most reliably measured (fluorescence intensity/background >3); using that data, our estimate of the percent variation in mRNA levels directly attributed to variation in gene copy number increases to 12% (Fig. 4*d*). This still undoubtedly represents a significant underestimate, as the observed variation in global gene expression is affected not only by true variation in the expression programs of the tumor cells themselves, but also by the variable presence of non-tumor cell types within clinical samples.

Discussion

This genome-wide, array CGH analysis of DNA copy number alteration in a series of human breast tumors demonstrates the usefulness of defining amplicon boundaries at high resolution (gene-by-gene), and quantitatively measuring amplicon shape, to assist in locating and identifying candidate oncogenes. By analyzing mRNA levels in parallel, we have also discovered that changes in DNA copy number have a large, pervasive, direct effect on global gene expression patterns in both breast cancer cell lines and tumors. Although the DNA microarrays used in our analysis may display a bias toward characterized and/or highly expressed genes, because we are examining such a large fraction of the genome (approximately 20% of all human genes), and because, as detailed above, we are likely underestimating the contribution of DNA copy number changes to altered gene expression, we believe our findings are likely to be generalizable (but would nevertheless still be remarkable if only applicable to this set of ~6,100 genes).

In budding yeast, aneuploidy has been shown to result in chromosome-wide gene expression biases (13). Two recent studies have begun to examine the global relationship between DNA copy number and gene expression in cancer cells. In agreement with our findings, Phillips *et al.* (14) have shown that with the acquisition of tumorigenicity in an immortalized prostate epithelial cell line, new chromosomal gains and losses resulted in a statistically significant respective increase and decrease in the average expression level of involved genes. In contrast, Platzer *et al.* (15) recently reported that in metastatic colon tumors only ~4% of genes within amplified regions were found more highly (>2-fold) expressed, when compared with normal colonic epithelium. This report differs substantially from our finding that 62% of highly amplified genes in breast cancer exhibit at least 2-fold increased expression. These contrasting findings may reflect methodological differences between the

GENETICS

studies. For example, the study of Platzer *et al.* (15) may have systematically under-measured gene expression changes. In this regard it is remarkable that only 14 transcripts of many thousand residing within unamplified chromosomal regions were found to exhibit at least 4-fold altered expression in metastatic colon cancer. Additionally, their reliance on lower-resolution chromosomal CGH may have resulted in poorly delimiting the boundaries of high-complexity amplicons, effectively overcalling regions with amplification. Alternatively, the contrasting findings for amplified genes may represent real biological differences between breast and metastatic colon tumors; resolution of this issue will require further studies.

Our finding that widespread DNA copy number alteration has a large, pervasive and direct effect on global gene expression patterns in breast cancer has several important implications. First, this finding supports a high degree of copy number-dependent gene expression in tumors. Second, it suggests that most genes are not subject to specific autoregulation or dosage compensation. Third, this finding cautions that elevated expression of an amplified gene cannot alone be considered strong independent evidence of a candidate oncogene's role in tumorigenesis. In our study, fully 62% of highly amplified genes demonstrated moderately or highly elevated expression. This highlights the importance of high-resolution mapping of amplicon boundaries and shape [to identify the "driving" gene(s) within amplicons (16)], on a large number of samples, in addition to functional studies. Fourth, this finding suggests that analyzing the genomic distribution of expressed genes, even within existing microarray gene expression data sets, may permit the inference of DNA copy number aberration, particularly aneuploidy (where gene expression can be averaged across large chromosomal regions; see Fig. 3 and supporting information). Fifth, this finding implies that a substantial portion of the phenotypic uniqueness (and by extension, the heterogeneity in clinical behavior) among patients' tumors may be traceable to underlying variation in DNA copy number. Sixth, this finding supports a possible role for widespread DNA copy number alteration in tumorigenesis (17, 18), beyond the amplification of specific oncogenes and deletion of specific tumor suppressor genes. Widespread DNA copy number alteration, and the concomitant widespread imbalance in gene expression, might disrupt critical stochiometric relationships in cell metabolism and physiology (e.g., proteosome, mitotic spindle), possibly promoting further chromosomal instability and directly contributing to tumor development or progression. Finally, our findings suggest the possibility of cancer therapies that exploit specific or global imbalances in gene expression in cancer.

We thank the many members of the P.O.B. and D.B. labs for helpful discussions. J.R.P. was a Howard Hughes Medical Institute Physician Postdoctoral Fellow during a portion of this work. P.O.B. is a Howard Hughes Medical Institute Associate Investigator. This work was supported by grants from the National Institutes of Health, the Howard Hughes Medical Institute, the Norwegian Cancer Society, and the Norwegian Research Council.

1. Kallioniemi, A., Kallioniemi, O. P., Sudar, D., Rutovitz, D., Gray, J. W., Waldman, F. & Pinkel, D. (1992) *Science* **258**, 818–821.
2. Kallioniemi, A., Kallioniemi, O. P., Piper, J., Tanner, M., Stokke, T., Chen, L., Smith, H. S., Pinkel, D., Gray, J. W. & Waldman, F. M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 2156–2160.
3. Tirkkonen, M., Tanner, M., Karhu, R., Kallioniemi, A., Isola, J. & Kallioniemi, O. P. (1998) *Genes Chromosomes Cancer* **21**, 177–184.
4. Forozan, F., Mahlamaki, E. H., Monni, O., Chen, Y., Veldman, R., Jiang, Y., Gooden, G. C., Ethier, S. P., Kallioniemi, A. & Kallioniemi, O. P. (2000) *Cancer Res.* **60**, 4519–4525.
5. Solinas-Toldo, S., Lampel, S., Stilgenbauer, S., Nickolenko, J., Benner, A., Dohner, H., Cremer, T. & Lichter, P. (1997) *Genes Chromosomes Cancer* **20**, 399–407.
6. Pinkel, D., Segraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., Collins, C., Kuo, W. L., Chen, C., Zhai, Y., *et al.* (1998) *Nat. Genet.* **20**, 207–211.
7. Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D. & Brown, P. O. (1999) *Nat. Genet.* **23**, 41–46.
8. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., *et al.* (2000) *Nature (London)* **406**, 747–752.
9. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **98**, 10869–10874.
10. Schuler, G. D. (1997) *J. Mol. Med.* **75**, 694–698.
11. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001) *Nature (London)* **409**, 860–921.
12. Fejzo, M. S., Godfrey, T., Chen, C., Waldman, F. & Gray, J. W. (1998) *Genes Chromosomes Cancer* **22**, 105–113.
13. Hughes, T. R., Roberts, C. J., Dai, H., Jones, A. R., Meyer, M. R., Slade, D., Burchard, J., Dow, S., Ward, T. R., Kidd, M. J., *et al.* (2000) *Nat. Genet.* **25**, 333–337.
14. Phillips, J. L., Hayward, S. W., Wang, Y., Vasselli, J., Pavlovich, C., Padilla-Nash, H., Pezullo, J. R., Ghadimi, B. M., Grossfeld, G. D., Rivera, A., *et al.* (2001) *Cancer Res.* **61**, 8143–8149.
15. Platzer, P., Upender, M. B., Wilson, K., Willis, J., Lutterbaugh, J., Nosrati, A., Willson, J. K., Mack, D., Ried, T. & Markowitz, S. (2002) *Cancer Res.* **62**, 1134–1138.
16. Albertson, D. G., Ylstra, B., Segraves, R., Collins, C., Dairkee, S. H., Kowbel, D., Kuo, W. L., Gray, J. W. & Pinkel, D. (2000) *Nat. Genet.* **25**, 144–146.
17. Li, R., Yerganian, G., Duesberg, P., Kraemer, A., Willer, A., Rausch, C. & Hehlmann, R. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 14506–14511.
18. Rasnick, D. & Duesberg, P. H. (1999) *Biochem. J.* **340**, 621–630.

Genome-wide Study of Gene Copy Numbers, Transcripts, and Protein Levels in Pairs of Non-invasive and Invasive Human Transitional Cell Carcinomas*

Torben F. Ørntoft‡§, Thomas Thykjaer¶, Frederic M. Waldman||, Hans Wolf**, and Julio E. Celis‡‡

Gain and loss of chromosomal material is characteristic of bladder cancer, as well as malignant transformation in general. The consequences of these changes at both the transcription and translation levels is at present unknown partly because of technical limitations. Here we have attempted to address this question in pairs of non-invasive and invasive human bladder tumors using a combination of technology that included comparative genomic hybridization, high density oligonucleotide array-based monitoring of transcript levels (5600 genes), and high resolution two-dimensional gel electrophoresis. The results showed that there is a gene dosage effect that in some cases superimposes on other regulatory mechanisms. This effect depended ($p < 0.015$) on the magnitude of the comparative genomic hybridization change. In general (18 of 23 cases), chromosomal areas with more than 2-fold gain of DNA showed a corresponding increase in mRNA transcripts. Areas with loss of DNA, on the other hand, showed either reduced or unaltered transcript levels. Because most proteins resolved by two-dimensional gels are unknown it was only possible to compare mRNA and protein alterations in relatively few cases of well focused abundant proteins. With few exceptions we found a good correlation ($p < 0.005$) between transcript alterations and protein levels. The implications, as well as limitations, of the approach are discussed. ***Molecular & Cellular Proteomics 1:37–45, 2002.***

Aneuploidy is a common feature of most human cancers (1), but little is known about the genome-wide effect of this phenomenon at both the transcription and translation levels. High throughput array studies of the breast cancer cell line BT474 has suggested that there is a correlation between DNA copy numbers and gene expression in highly amplified areas (2), and studies of individual genes in solid tumors have revealed a good correlation between gene dose and mRNA or protein levels in the case of c-erb-B2, *cyclin d1*, *ems1*, and N-myc (3–5). However, a high cyclin D1 protein expression has been observed without simultaneous amplification (4), and a low level of c-myc copy number increase was observed without concomitant c-myc protein overexpression (6).

In human bladder tumors, karyotyping, fluorescent *in situ* hybridization, and comparative genomic hybridization (CGH)[1] have revealed chromosomal aberrations that seem to be characteristic of certain stages of disease progression. In the case of non-invasive pTa transitional cell carcinomas (TCCs), this includes loss of chromosome 9 or parts of it, as well as loss of Y in males. In minimally invasive pT1 TCCs, the following alterations have been reported: 2q−, 11p−, 1q+, 11q13+, 17q+, and 20q+ (7–12). It has been suggested that these regions harbor tumor suppressor genes and oncogenes; however, the large chromosomal areas involved often contain many genes, making meaningful predictions of the functional consequences of losses and gains very difficult.

In this investigation we have combined genome-wide technology for detecting genomic gains and losses (CGH) with gene expression profiling techniques (microarrays and proteomics) to determine the effect of gene copy number on transcript and protein levels in pairs of non-invasive and invasive human bladder TCCs.

EXPERIMENTAL PROCEDURES

Material—Bladder tumor biopsies were sampled after informed consent was obtained and after removal of tissue for routine pathology examination. By light microscopy tumors 335 and 532 were staged by an experienced pathologist as pTa (superficial papillary),

From the ‡Department of Clinical Biochemistry, Molecular Diagnostic Laboratory and **Department of Urology, Aarhus University Hospital, Skejby, DK-8200 Aarhus N, Denmark, ¶AROS Applied Biotechnology ApS, Gustav Wiedsvej 10, DK-8000 Aarhus C, Denmark, ||UCSF Cancer Center and Department of Laboratory Medicine, University of California, San Francisco, CA 94143-0808, and ‡‡Institute of Medical Biochemistry and Danish Centre for Human Genome Research, Ole Worms Allé 170, Aarhus University, DK-8000 Aarhus C, Denmark

Received, September 26, 2001, and in revised form, November 7, 2001

Published, MCP Papers in Press, November 13, 2001, DOI 10.1074/mcp.M100019-MCP200

[1] The abbreviations used are: CGH, comparative genomic hybridization; TCC, transitional cell carcinoma; LOH, loss of heterozygosity; PA-FABP, psoriasis-associated fatty acid-binding protein; 2D, two-dimensional.

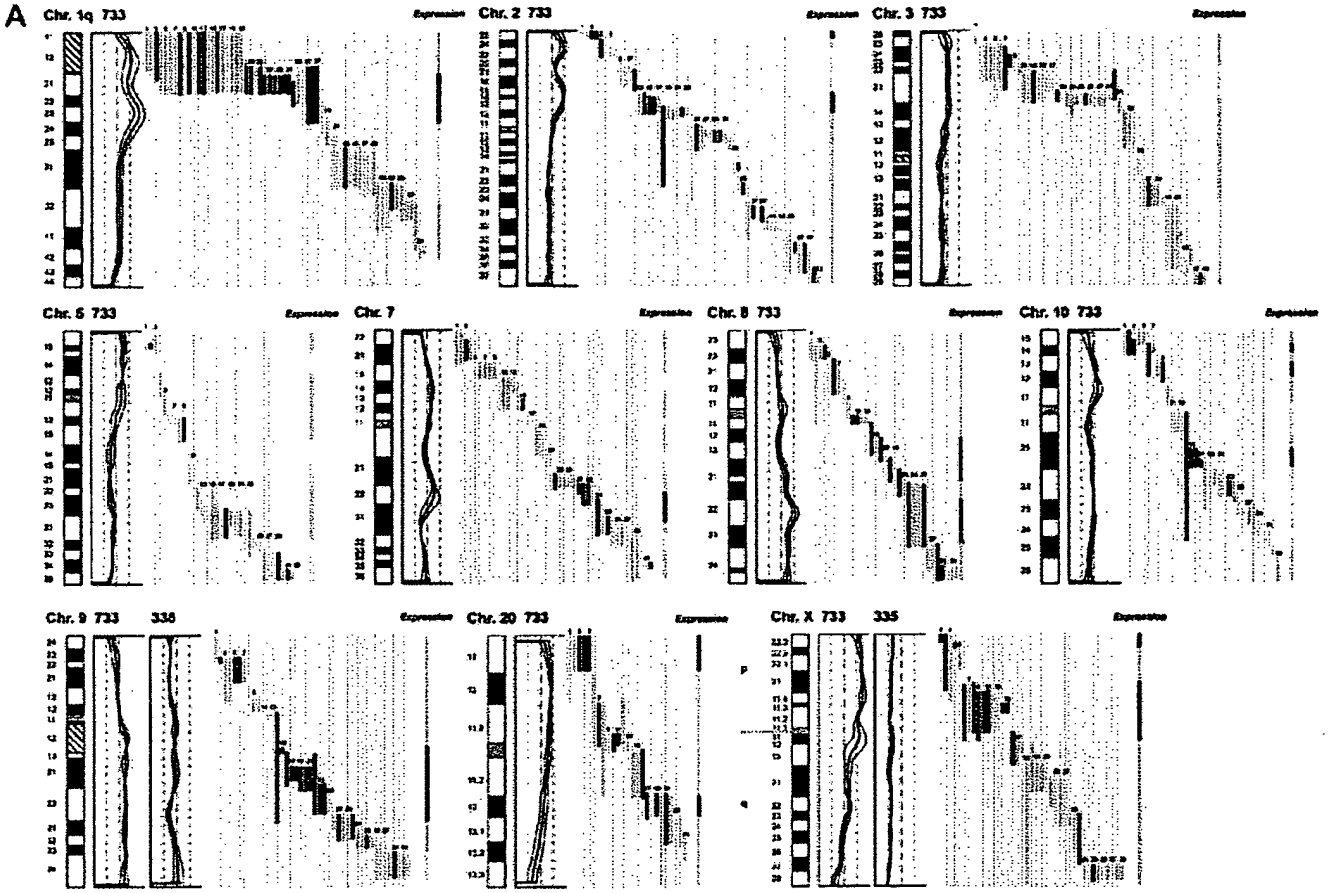


FIG. 1. **DNA copy number and mRNA expression level.** Shown from *left* to *right* are chromosome (*Chr.*), CGH profiles, gene location and expression level of specific genes, and overall expression level along the chromosome. *A*, expression of mRNA in invasive tumor 733 as compared with the non-invasive counterpart tumor 335. *B*, expression of mRNA in invasive tumor 827 compared with the non-invasive counterpart tumor 532. The average fluorescent signal ratio between tumor DNA and normal DNA is shown along the length of the chromosome (*left*). The *bold curve* in the ratio profile represents a mean of four chromosomes and is surrounded by *thin curves* indicating one standard deviation. The *central vertical line* (*broken*) indicates a ratio value of 1 (no change), and the *vertical lines* next to it (*dotted*) indicate a ratio of 0.5 (*left*) and 2.0 (*right*). In chromosomes where the non-invasive tumor 335 used for comparison showed alterations in DNA content, the ratio profile of that chromosome is shown to the *right* of the invasive tumor profile. The *colored bars* represents one gene each, identified by the running *numbers above* the *bars* (the name of the gene can be seen at www.MDL.DK/sdata.html). The *bars* indicate the purported location of the gene, and the *colors* indicate the expression level of the gene in the invasive tumor compared with the non-invasive counterpart; >2-fold increase (*black*), >2-fold decrease (*blue*), no significant change (*orange*). The *bar* to the *far right*, entitled *Expression* shows the resulting change in expression along the chromosome; the *colors* indicate that at least half of the genes were up-regulated (*black*), at least half of the genes down-regulated (*blue*), or more than half of the genes are unchanged (*orange*). If a gene was absent in one of the samples and present in another, it was regarded as more than a 2-fold change. A 2-fold level was chosen as this corresponded to one standard deviation in a double determination of ~1800 genes. Centromeres and heterochromatic regions were excluded from data analysis.

grade I and II, respectively, tumors 733 and 827 were staged as pT1 (invasive into submucosa), 733 was staged as solid, and 827 was staged as papillary, both grade III.

mRNA Preparation—Tissue biopsies, obtained fresh from surgery, were embedded immediately in a sodium-guanidinium thiocyanate solution and stored at −80 °C. Total RNA was isolated using the RNAzol B RNA isolation method (WAK-Chemie Medical GMBH). poly(A)$^+$ RNA was isolated by an oligo(dT) selection step (Oligotex mRNA kit; Qiagen).

cRNA Preparation—1 μg of mRNA was used as starting material. The first and second strand cDNA synthesis was performed using the SuperScript® choice system (Invitrogen) according to the manufacturer's instructions but using an oligo(dT) primer containing a T7 RNA polymerase binding site. Labeled cRNA was prepared using the MEGAscrip® *in vitro* transcription kit (Ambion). Biotin-labeled CTP and UTP (Enzo) was used, together with unlabeled NTPs in the reaction. Following the *in vitro* transcription reaction, the unincorporated nucleotides were removed using RNeasy columns (Qiagen).

Array Hybridization and Scanning—Array hybridization and scanning was modified from a previous method (13). 10 μg of cRNA was fragmented at 94 °C for 35 min in buffer containing 40 mM Tris acetate, pH 8.1, 100 mM KOAc, 30 mM MgOAc. Prior to hybridization, the fragmented cRNA in a 6× SSPE-T hybridization buffer (1 M NaCl, 10 mM Tris, pH 7.6, 0.005% Triton), was heated to 95 °C for 5 min, subsequently cooled to 40 °C, and loaded onto the Affymetrix probe array cartridge. The probe array was then incubated for 16 h at 40 °C at constant rotation (60 rpm). The probe array was exposed to 10 washes in 6× SSPE-T at 25 °C followed by 4 washes in 0.5× SSPE-T at 50 °C. The biotinylated cRNA was stained with a streptavidin-phycoerythrin conjugate, 10 μg/ml (Molecular Probes) in 6× SSPE-T

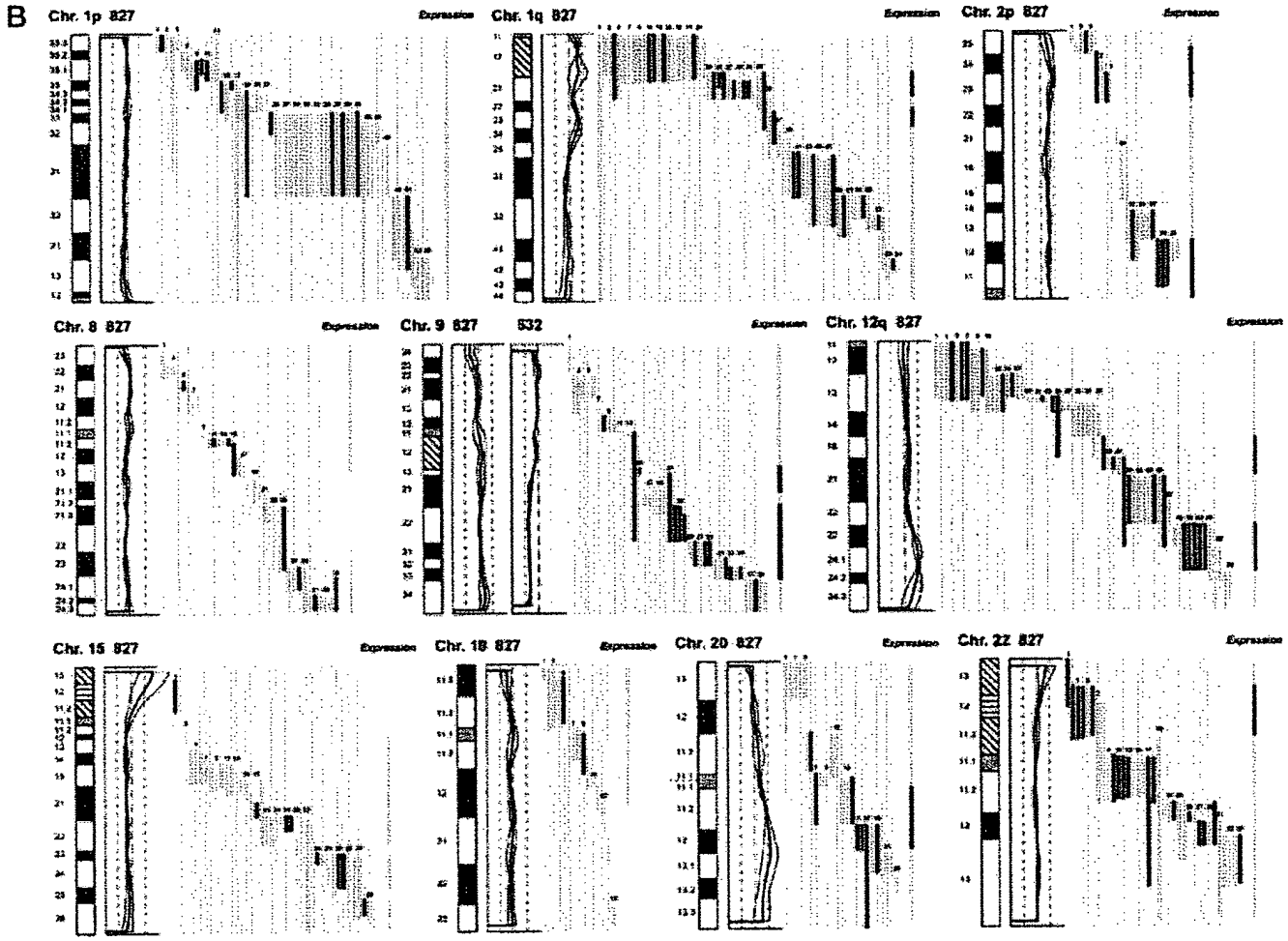


FIG. 1—*continued*

for 30 min at 25 °C followed by 10 washes in 6× SSPE-T at 25 °C. The probe arrays were scanned at 560 nm using a confocal laser scanning microscope (made for Affymetrix by Hewlett-Packard). The readings from the quantitative scanning were analyzed by Affymetrix gene expression analysis software.

Microsatellite Analysis—Microsatellite Analysis was performed as described previously (14). Microsatellites were selected by use of www.ncbi.nlm.nih.gov/genemap98, and primer sequences were obtained from the genome data base at www.gdb.org. DNA was extracted from tumor and blood and amplified by PCR in a volume of 20 μl for 35 cycles. The amplicons were denatured and electrophoresed for 3 h in an ABI Prism 377. Data were collected in the Gene Scan program for fragment analysis. Loss of heterozygosity was defined as less than 33% of one allele detected in tumor amplicons compared with blood.

Proteomic Analysis—TCCs were minced into small pieces and homogenized in a small glass homogenizer in 0.5 ml of lysis solution. Samples were stored at −20 °C until use. The procedure for 2D gel electrophoresis has been described in detail elsewhere (15, 16). Gels were stained with silver nitrate and/or Coomassie Brilliant Blue. Proteins were identified by a combination of procedures that included microsequencing, mass spectrometry, two-dimensional gel Western immunoblotting, and comparison with the master two-dimensional gel image of human keratinocyte proteins; see biobase.dk/cgi-bin/celis.

CGH—Hybridization of differentially labeled tumor and normal DNA to normal metaphase chromosomes was performed as described previously (10). Fluorescein-labeled tumor DNA (200 ng), Texas Red-labeled reference DNA (200 ng), and human Cot-1 DNA (20 μg) were denatured at 37 °C for 5 min and applied to denatured normal metaphase slides. Hybridization was at 37 °C for 2 days. After washing, the slides were counterstained with 0.15 μg/ml 4,6-diamidino-2-phenylindole in an anti-fade solution. A second hybridization was performed for all tumor samples using fluorescein-labeled reference DNA and Texas Red-labeled tumor DNA (inverse labeling) to confirm the aberrations detected during the initial hybridization. Each CGH experiment also included a normal control hybridization using fluorescein- and Texas Red-labeled normal DNA. Digital image analysis was used to identify chromosomal regions with abnormal fluorescence ratios, indicating regions of DNA gains and losses. The average green:red fluorescence intensity ratio profiles were calculated using four images of each chromosome (eight chromosomes total) with normalization of the green:red fluorescence intensity ratio for the entire metaphase and background correction. Chromosome identification was performed based on 4,6-diamidino-2-phenylindole banding patterns. Only images showing uniform high intensity fluorescence with minimal background staining were analyzed. All centromeres, p arms of acrocentric chromosomes, and heterochromatic regions were excluded from the analysis.

RESULTS

Comparative Genomic Hybridization—The CGH analysis identified a number of chromosomal gains and losses in the

TABLE I

Correlation between alterations detected by CGH and by expression monitoring

Top, CGH used as independent variable (if CGH alteration – what expression ratio was found); bottom, altered expression used as independent variable (if expression alteration – what CGH deviation was found).

CGH alterations	Tumor 733 vs. 335 Expression change clusters	Concordance	CGH alterations	Tumor 827 vs. 532 Expression change clusters	Concordance
13 Gain	10 Up-regulation 0 Down-regulation 3 No change	77%	10 Gain	8 Up-regulation 0 Down-regulation 2 No change	80%
10 Loss	1 Up-regulation 5 Down-regulation 4 No change	50%	12 Loss	3 Up-regulation 2 Down regulation 7 No change	17%
Expression change clusters	**Tumor 733 vs. 335 CGH alterations**	**Concordance**	**Expression change clusters**	**Tumor 827 vs. 532 CGH alterations**	**Concordance**
16 Up-regulation	11 Gain 2 Loss 3 No change	69%	17 Up-regulation	10 Gain 5 Loss 2 No change	59%
21 Down-regulation	1 Gain 8 Loss 12 No change	38%	9 Down-regulation	0 Gain 3 Loss 6 No change	33%
15 No change	3 Gain 3 Loss 9 No change	60%	21 No change	1 Gain 3 Loss 17 No change	81%

two invasive tumors (stage pT1, TCCs 733 and 827), whereas the two non-invasive papillomas (stage pTa, TCCs 335 and 532) showed only 9p–, 9q22-q33–, and X–, and 7+, 9q–, and Y–, respectively. Both invasive tumors showed changes (1q22-24+, 2q14.1-qter–, 3q12-q13.3–, 6q12-q22–, 9q34+, 11q12-q13+, 17+, and 20q11.2-q12+) that are typical for their disease stage, as well as additional alterations, some of which are shown in Fig. 1. Areas with gains and losses deviated from the normal copy number to some extent, and the average numerical deviation from normal was 0.4-fold in the case of TCC 733 and 0.3-fold for TCC 827. The largest changes, amounting to at least a doubling of chromosomal content, were observed at 1q23 in TCC 733 (Fig. 1*A*) and 20q12 in TCC 827 (Fig. 1*B*).

mRNA Expression in Relation to DNA Copy Number—The mRNA levels from the two invasive tumors (TCCs 827 and 733) were compared with the two non-invasive counterparts (TCCs 532 and 335). This was done in two separate experiments in which we compared TCCs 733 to 335 and 827 to 532, respectively, using two different scaling settings for the arrays to rule out scaling as a confounding parameter. Approximately 1,800 genes that yielded a signal on the arrays were searched in the Unigene and Genemap data bases for chromosomal location, and those with a known location (1096) were plotted as bars covering their purported locus. In that way it was possible to construct a graphic presentation of DNA copy number and relative mRNA levels along the individual chromosomes (Fig. 1).

For each mRNA a ratio was calculated between the level in the invasive *versus* the non-invasive counterpart. Bars, which represent chromosomal location of a gene, were color-coded according to the expression ratio, and only differences larger than 2-fold were regarded as informative (Fig. 1). The density of genes along the chromosomes varied, and areas containing only one gene were excluded from the calculations. The resolution of the CGH method is very low, and some of the outlier data may be because of the fact that the boundaries of the chromosomal aberrations are not known at high resolution.

Two sets of calculations were made from the data. For the first set we used CGH alterations as the independent variable and estimated the frequency of expression alterations in these chromosomal areas. In general, areas with a strong gain of chromosomal material contained a cluster of genes having increased mRNA expression. For example, both chromosomes 1q21-q25, 2p and 9q, showed a relative gain of more than 100% in DNA copy number that was accompanied by increased mRNA expression levels in the two tumor pairs (Fig. 1). In most cases, chromosomal gains detected by CGH were accompanied by an increased level of transcripts in both TCCs 733 (77%) and 827 (80%) (Table I, *top*). Chromosomal losses, on the other hand, were not accompanied by decreased expression in several cases, and were often registered as having unaltered RNA levels (Table I, *top*). The inability to detect RNA expression changes in these cases was not because of fewer genes mapping to the lost regions (data not shown).

In the second set of calculations we selected expression alterations above 2-fold as the independent variable and estimated the frequency of CGH alterations in these areas. As above, we found that increased transcript expression correlated with gain of chromosomal material (TCC 733, 69% and TCC 827, 59%), whereas reduced expression was often detected in areas with unaltered CGH ratios (Table I, *bottom*). Furthermore, as a control we looked at areas with no alter-

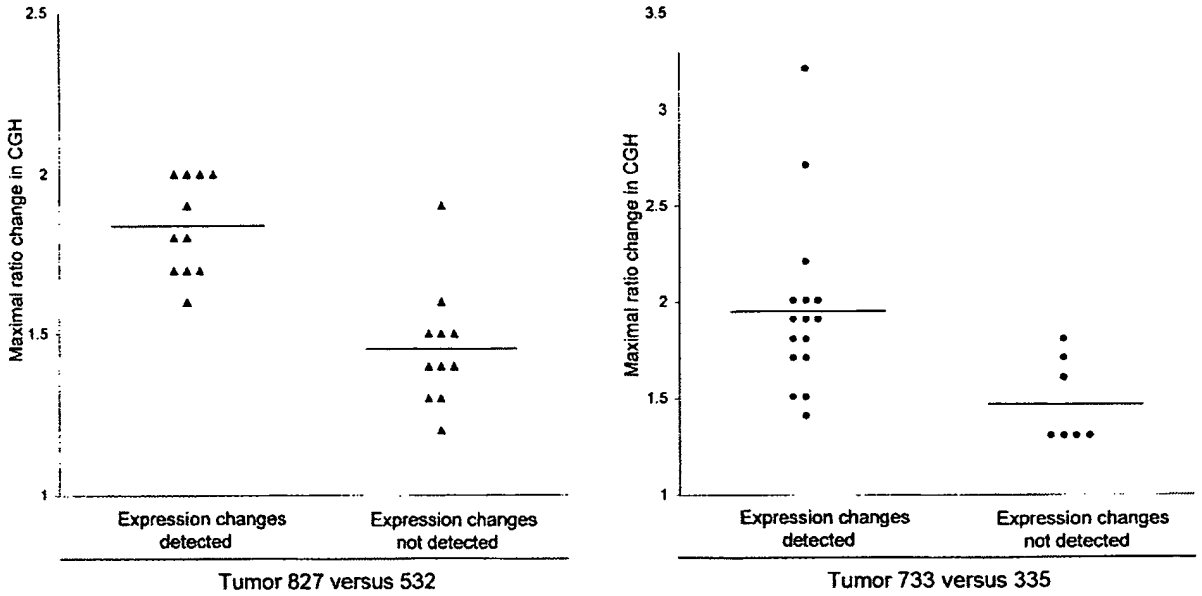


FIG. 2. **Correlation between maximum CGH aberration and the ability to detect expression change by oligonucleotide array monitoring.** The aberration is shown as a numerical -fold change in ratio between invasive tumors 827 (▲) and 733 (♦) and their non-invasive counterparts 532 and 335. The expression change was taken from the *Expression* line to the *right* in Fig. 1, which depicts the resulting expression change for a given chromosomal region. At least half of the mRNAs from a given region have to be either up- or down-regulated to be scored as an expression change. All chromosomal arms in which the CGH ratio plus or minus one standard deviation was outside the ratio value of one were included.

ation in expression. No alteration was detected by CGH in most of these areas (TCC 733, 60% and TCC 827, 81%; see Table I, *bottom*). Because the ability to observe reduced or increased mRNA expression clustering to a certain chromosomal area clearly reflected the extent of copy number changes, we plotted the maximum CGH aberrations in the regions showing CGH changes against the ability to detect a change in mRNA expression as monitored by the oligonucleotide arrays (Fig. 2). For both tumors TCC 733 ($p < 0.015$) and TCC 827 ($p < 0.00003$) a highly significant correlation was observed between the level of CGH ratio change (reflecting the DNA copy number) and alterations detected by the array based technology (Fig. 2). Similar data were obtained when areas with altered expression were used as independent variables. These areas correlated best with CGH when the CGH ratio deviated 1.6- to 2.0-fold (Table I, *bottom*) but mostly did not at lower CGH deviations. These data probably reflect that loss of an allele may only lead to a 50% reduction in expression level, which is at the cut-off point for detection of expression alterations. Gain of chromosomal material can occur to a much larger extent.

Microsatellite-based Detection of Minor Areas of Losses—In TCC 733, several chromosomal areas exhibiting DNA amplification were preceded or followed by areas with a normal CGH but reduced mRNA expression (see Fig. 1, TCC 733 chromosome 1q32, 2p21, and 7q21 and q32, 9q34, and 10q22). To determine whether these results were because of undetected loss of chromosomal material in these regions or because of other non-structural mechanisms regulating transcription, we examined two microsatellites positioned at chromosome 1q25–32 and two at chromosome 2p22. Loss of heterozygosity (LOH) was found at both 1q25 and at 2p22 indicating that minor deleted areas were not detected with the resolution of CGH (Fig. 3). Additionally, chromosome 2p in TCC 733 showed a CGH pattern of gain/no change/gain of DNA that correlated with transcript increase/decrease/increase. Thus, for the areas showing increased expression there was a correlation with the DNA copy number alterations (Fig. 1*A*). As indicated above, the mRNA decrease observed in the middle of the chromosomal gain was because of LOH, implying that one of the mechanisms for mRNA down-regulation may be regions that have undergone smaller losses of chromosomal material. However, this cannot be detected with the resolution of the CGH method.

In both TCC 733 and TCC 827, the telomeric end of chromosome 11p showed a normal ratio in the CGH analysis; however, clusters of five and three genes, respectively, lost their expression. Two microsatellites (D11S1760, D11S922) positioned close to MUC2, IGF2, and cathepsin D indicated LOH as the most likely mechanism behind the loss of expression (data not shown).

A reduced expression of mRNA observed in TCC 733 at chromosomes 3q24, 11p11, 12p12.2, 12q21.1, and 16q24 and in TCC 827 at chromosome 11p15.5, 12p11, 15q11.2, and 18q12 was also examined for chromosomal losses using microsatellites positioned as close as possible to the gene loci

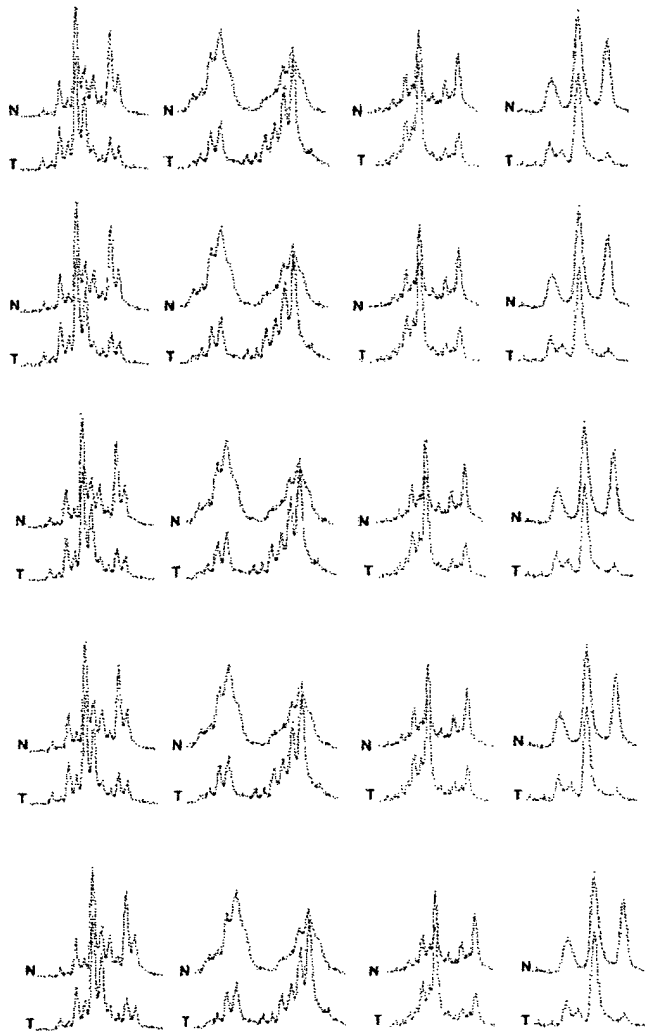


FIG. 3. **Microsatellite analysis of loss of heterozygosity.** Tumor 733 showing loss of heterozygosity at chromosome 1q25, detected (*a*) by D1S215 close to Hu class I histocompatibility antigen (gene number 38 in Fig. 1), (*b*) by D1S2735 close to cathepsin E (gene number 41 in Fig. 1), and (*c*) at chromosome 2p23 by D2S2251 close to general β-spectrin (gene number 11 on Fig. 1) and of (*d*) tumor 827 showing loss of heterozygosity at chromosome 18q12 by S18S1118 close to mitochondrial 3-oxoacyl-coenzyme A thiolase (gene number 12 in Fig. 1). The *upper curves* show the electropherogram obtained from normal DNA from leukocytes (*N*), and the *lower curves* show the electropherogram from tumor DNA (*T*). In all cases one allele is partially lost in the tumor amplicon.

showing reduced mRNA transcripts. Only the microsatellite positioned at 18q12 showed LOH (Fig. 3), suggesting that transcriptional down-regulation of genes in the other regions may be controlled by other mechanisms.

Relation between Changes in mRNA and Protein Levels—2D-PAGE analysis, in combination with Coomassie Brilliant Blue and/or silver staining, was carried out on all four tumors using fresh biopsy material. 40 well resolved abundant known proteins migrating in areas away from the edges of the pH

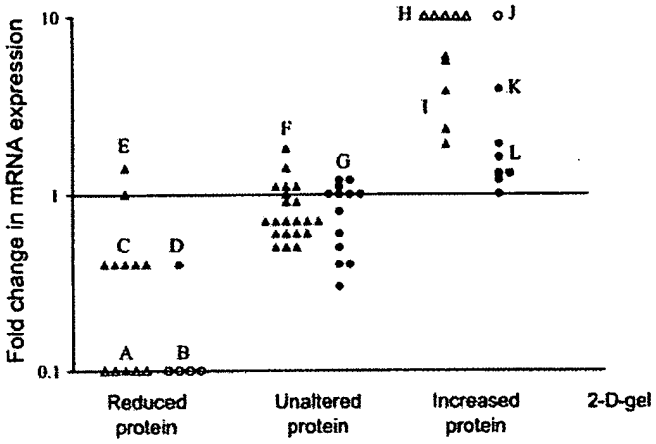


FIG. 4. **Correlation between protein levels as judged by 2D-PAGE and transcript ratio.** For comparison proteins were divided in three groups, unaltered in level or up- or down-regulated (*horizontal axis*). The mRNA ratio as determined by oligonucleotide arrays was plotted for each gene (*vertical axis*). ▲, mRNAs that were scored as present in both tumors used for the ratio calculation; △, mRNAs that were scored as absent in the invasive tumors (along *horizontal axis*) or as absent in non-invasive reference (*top* of figure). Two different scalings were used to exclude scaling as a confounder, TCCs 827 and 532 (▲△) were scaled with background suppression, and TCCs 733 and 335 (●○) were scaled without suppression. Both comparisons showed highly significant ($p < 0.005$) differences in mRNA ratios between the groups. Proteins shown were as follows: Group *A* (from *left*), phosphoglucomutase 1, glutathione transferase class μ number 4, fatty acid-binding protein homologue, cytokeratin 15, and cytokeratin 13; *B* (from *left*), fatty acid-binding protein homologue, 28-kDa heat shock protein, cytokeratin 13, and calcyclin; *C* (from *left*), α-enolase, hnRNP B1, 28-kDa heat shock protein, 14-3-3-ε, and pre-mRNA splicing factor; *D*, mesothelial keratin K7 (type II); *E* (from *top*), glutathione *S*-transferase-π and mesothelial keratin K7 (type II); *F* (from *top* and *left*), adenylyl cyclase-associated protein, E-cadherin, keratin 19, calgizzarin, phosphoglycerate mutase, annexin IV, cytoskeletal γ-actin, hnRNP A1, integral membrane protein calnexin (IP90), hnRNP H, brain-type clathrin light chain-a, hnRNP F, 70-kDa heat shock protein, heterogeneous nuclear ribonucleoprotein A/B, translationally controlled tumor protein, liver glyceraldehyde-3-phosphate dehydrogenase, keratin 8, aldehyde reductase, and Na,K-ATPase β-1 subunit; *G*, (from *top* and *left*), TCP20, calgizzarin, 70-kDa heat shock protein, calnexin, hnRNP H, cytokeratin 15, ATP synthase, keratin 19, triosephosphate isomerase, hnRNP F, liver glyceraldehyde-3-phosphatase dehydrogenase, glutathione *S*-transferase-π, and keratin 8; *H* (from *left*), plasma gelsolin, autoantigen calreticulin, thioredoxin, and NAD+-dependent 15 hydroxyprostaglandin dehydrogenase; *I* (from *top*), prolyl 4-hydroxylase β-subunit, cytokeratin 20, cytokeratin 17, prohibition, and fructose 1,6-biphosphatase; *J* annexin II; *K*, annexin IV; *L* (from *top* and *left*), 90-kDa heat shock protein, prolyl 4-hydroxylase β-subunit, α-enolase, GRP 78, cyclophilin, and cofilin.

gradient, and having a known chromosomal location, were selected for analysis in the TCC pair 827/532. Proteins were identified by a combination of methods (see "Experimental Procedures"). In general there was a highly significant correlation ($p < 0.005$) between mRNA and protein alterations (Fig. 4). Only one gene showed disagreement between transcript alteration and protein alteration. Except for a group of cyto-.

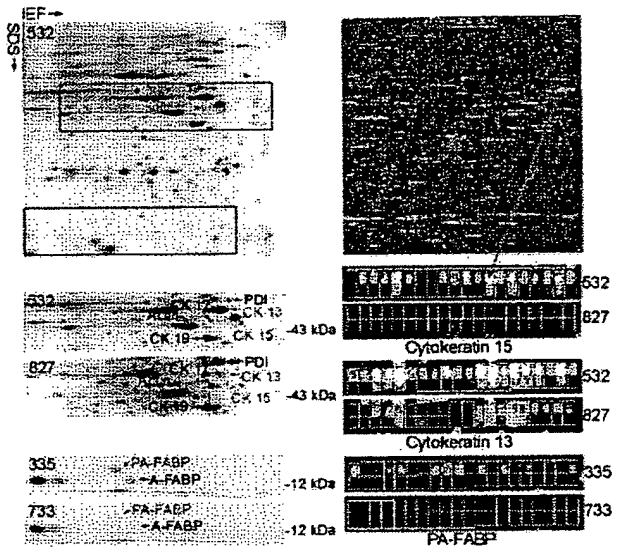


FIG. 5. **Comparison of protein and transcript levels in invasive and non-invasive TCCs.** The *upper part* of the figure shows a 2D gel (*left*) and the oligonucleotide array (*right*) of TCC 532. The *red rectangles* on the *upper gel* highlight the areas that are compared below. Identical areas of 2D gels of TCCs 532 and 827 are shown below. Clearly, cytokeratins 13 and 15 are strongly down-regulated in TCC 827 (*red annotation*). The tile on the array containing probes for cytokeratin 15 is enlarged *below* the array (*red arrow*) from TCC 532 and is compared with TCC 827. The *upper row* of *squares* in each tile corresponds to perfect match probes; the *lower row* corresponds to mismatch probes containing a mutation (used for correction for unspecific binding). Absence of signal is depicted as *black*, and the higher the signal the lighter the color. A high transcript level was detected in TCC 532 (6151 units) whereas a much lower level was detected in TCC 827 (absence of signals). For cytokeratin 13, a high transcript level was also present in TCC 532 (15659 units), and a much lower level was present in TCC 827 (623 units). The 2D gels at the *bottom* of the figure (*left*) show levels of PA-FABP and adipocyte-FABP in TCCs 335 and 733 (invasive), respectively. Both proteins are down-regulated in the invasive tumor. To the *right* we show the array tiles for the PA-FABP transcript. A medium transcript level was detected in the case of TCC 335 (1277 units) whereas very low levels were detected in TCC 733 (166 units). *IEF*, isoelectric focusing.

keratins encoded by genes on chromosome 17 (Fig. 5) the analyzed proteins did not belong to a particular family. 26 well focused proteins whose genes had a know chromosomal location were detected in TCCs 733 and 335, and of these 19 correlated ($p < 0.005$) with the mRNA changes detected using the arrays (Fig. 4). For example, PA-FABP was highly expressed in the non-invasive TCC 335 but lost in the invasive counterpart (TCC 733; see Fig. 5). The smaller number of proteins detected in both 733 and 335 was because of the smaller size of the biopsies that were available.

11 chromosomal regions where CGH showed aberrations that corresponded to the changes in transcript levels also showed corresponding changes in the protein level (Table II). These regions included genes that encode proteins that are found to be frequently altered in bladder cancer, namely cytokeratins 17 and 20, annexins II and IV, and the fatty acid-binding proteins PA-FABP and FBP1. Four of these proteins were encoded by genes in chromosome 17q, a frequently amplified chromosomal area in invasive bladder cancers.

DISCUSSION

Most human cancers have abnormal DNA content, having lost some chromosomal parts and gained others. The present study provides some evidence as to the effect of these gains and losses on gene expression in two pairs of non-invasive and invasive TCCs using high throughput expression arrays and proteomics, in combination with CGH. In general, the results showed that there is a clear individual regulation of the mRNA expression of single genes, which in some cases was superimposed by a DNA copy number effect. In most cases, genes located in chromosomal areas with gains often exhibited increased mRNA expression, whereas areas showing losses showed either no change or a reduced mRNA expression. The latter might be because of the fact that losses most often are restricted to loss of one allele, and the cut-off point for detection of expression alterations was a 2-fold change, thus being at the border of detection. In several cases, how-

TABLE II

Proteins whose expression level correlates with both mRNA and gene dose changes

Protein	Chromosomal location	Tumor TCC	CGH alteration	Transcript alteration[a]	Protein alteration
Annexin II	1q21	733	Gain	Abs to Pres[a]	Increase
Annexin IV	2p13	733	Gain	3.9-Fold up	Increase
Cytokeratin 17	17q12-q21	827	Gain	3.8-Fold up	Increase
Cytokeratin 20	17q21.1	827	Gain	5.6-Fold up	Increase
(PA-)FABP	8q21.2	827	Loss	10-Fold down	Decrease
FBP1	9q22	827	Gain	2.3-Fold up	Increase
Plasma gelsolin	9q31	827	Gain	Abs to Pres	Increase
Heat shock protein 28	15q12-q13	827	Loss	2.5-Fold up	Decrease
Prohibitin	17q21	827/733	Gain	3.7-/2.5-Fold up[b]	Increase
Prolyl-4-hydroxyl	17q25	827/733	Gain	5.7-/1.6-Fold up	Increase
hnRNPB1	7p15	827	Loss	2.5-Fold down	Decrease

[a] Abs, absent; Pres, present.
[b] In cases where the corresponding alterations were found in both TCCs 827 and 733 these are shown as 827/733.

ever, an increase or decrease in DNA copy number was associated with *de novo* occurrence or complete loss of transcript, respectively. Some of these transcripts could not be detected in the non-invasive tumor but were present at relatively high levels in areas with DNA amplifications in the invasive tumors (*e.g.* in TCC 733 transcript from cellular ligand of annexin II gene (chromosome 1q21) from absent to 2670 arbitrary units; in TCC 827 transcript from small proline-rich protein 1 gene (chromosome 1q12-q21.1) from absent to 1326 arbitrary units). It may be anticipated from these data that significant clustering of genes with an increased expression to a certain chromosomal area indicates an increased likelihood of gain of chromosomal material in this area.

Considering the many possible regulatory mechanisms acting at the level of transcription, it seems striking that the gene dose effects were so clearly detectable in gained areas. One hypothetical explanation may lie in the loss of controlled methylation in tumor cells (17–19). Thus, it may be possible that in chromosomes with increased DNA copy numbers two or more alleles could be demethylated simultaneously leading to a higher transcription level, whereas in chromosomes with losses the remaining allele could be partly methylated, turning off the process (20, 21). A recent report has documented a ploidy regulation of gene expression in yeast, but in this case all the genes were present in the same ratio (22), a situation that is not analogous to that of cancer cells, which show marked chromosomal aberrations, as well as gene dosage effects.

Several CGH studies of bladder cancer have shown that some chromosomal aberrations are common at certain stages of disease progression, often occurring in more than 1 of 3 tumors. In pTa tumors, these include 9p−, 9q−, 1q+, Y− (2, 6), and in pT1 tumors, 2q−,11p−, 11q−, 1q+, 5p+, 8q+, 17q+, and 20q+ (2–4, 6, 7). The pTa tumors studied here showed similar aberrations such as 9p− and 9q22-q33− and 9q− and Y−, respectively. Likewise, the two minimal invasive pT1 tumors showed aberrations that are commonly seen at that stage, and TCC 827 had a remarkable resemblance to the commonly seen pattern of losses and gains, such as 1q22–24 amplification (seen in both tumors), 11q14-q22 loss, the latter often linked to 17 q+ (both tumors), and 1q+ and 9p−, often linked to 20q+ and 11 q13+ (both tumors) (7–9). These observations indicate that the pairs of tumors used in this study exhibit chromosomal changes observed in many tumors, and therefore the findings could be of general importance for bladder cancer.

Considering that the mapping resolution of CGH is of about 20 megabases it is only possible to get a crude picture of chromosomal instability using this technique. Occasionally, we observed reduced transcript levels close to or inside regions with increased copy numbers. Analysis of these regions by positioning heterozygous microsatellites as close as possible to the locus showing reduced gene expression revealed loss of heterozygosity in several cases. It seems likely that multiple and different events occur along each chromosomal arm and that the use of cDNA microarrays for analysis of DNA copy number changes will reach a resolution that can resolve these changes, as has recently been proposed (2). The outlier data were not more frequent at the boundaries of the CGH aberrations. At present we do not know the mechanism behind chromosomal aneuploidy and cannot predict whether chromosomal gains will be transcribed to a larger extent than the two native alleles. A mechanism as genetic imprinting has an impact on the expression level in normal cells and is often reduced in tumors. However, the relation between imprinting and gain of chromosomal material is not known.

We regard it as a strength of this investigation that we were able to compare invasive tumors to benign tumors rather than to normal urothelium, as the tumors studied were biologically very close and probably may represent successive steps in the progression of bladder cancer. Despite the limited amount of fresh tissue available it was possible to apply three different state of the art methods. The observed correlation between DNA copy number and mRNA expression is remarkable when one considers that different pieces of the tumor biopsies were used for the different sets of experiments. This indicate that bladder tumors are relatively homogenous, a notion recently supported by CGH and LOH data that showed a remarkable similarity even between tumors and distant metastasis (10, 23).

In the few cases analyzed, mRNA and protein levels showed a striking correspondence although in some cases we found discrepancies that may be attributed to translational regulation, post-translational processing, protein degradation, or a combination of these. Some transcripts belong to undertranslated mRNA pools, which are associated with few translationally inactive ribosomes; these pools, however, seem to be rare (24). Protein degradation, for example, may be very important in the case of polypeptides with a short half-life (*e.g.* signaling proteins). A poor correlation between mRNA and protein levels was found in liver cells as determined by arrays and 2D-PAGE (25), and a moderate correlation was recently reported by Ideker *et al.* (26) in yeast.

Interestingly, our study revealed a much better correlation between gained chromosomal areas and increased mRNA levels than between loss of chromosomal areas and reduced mRNA levels. In general, the level of CGH change determined the ability to detect a change in transcript. One possible explanation could be that by losing one allele the change in mRNA level is not so dramatic as compared with gain of material, which can be rather unlimited and may lead to a severalfold increase in gene copy number resulting in a much higher impact on transcript level. The latter would be much easier to detect on the expression arrays as the cut-off point was placed at a 2-fold level so as not to be biased by noise on the array. Construction of arrays with a better signal to noise ratio may in the future allow detection of lesser than 2-fold alterations in transcript levels, a feature that may facilitate the analysis of the effect of loss of chromosomal areas on transcript levels.

In eleven cases we found a significant correlation between DNA copy number, mRNA expression, and protein level. Four of these proteins were encoded by genes located at a frequently amplified area in chromosome 17q. Whether DNA copy number is one of the mechanisms behind alteration of these eleven proteins is at present unknown and will have to be proved by other methods using a larger number of samples. One factor making such studies complicated is the large extent of protein modification that occurs after translation, requiring immunoidentification and/or mass spectrometry to correctly identify the proteins in the gels.

In conclusion, the results presented in this study exemplify the large body of knowledge that may be possible to gather in the future by combining state of the art techniques that follow the pathway from DNA to protein (26). Here, we used a traditional chromosomal CGH method, but in the future high resolution CGH based on microarrays with many thousand radiation hybrid-mapped genes will increase the resolution and information derived from these types of experiments (2). Combined with expression arrays analyzing transcripts derived from genes with known locations, and 2D gel analysis to obtain information at the post-translational level, a clearer and more developed understanding of the tumor genome will be forthcoming.

Acknowledgments—We thank Mie Madsen, Hanne Steen, Inge Lis Thorsen, Hans Lund, Vikolaj Ørntoft, and Lynn Bjerke for technical help and Thomas Gingeras, Christine Harrington, and Morten Østergaard for valuable discussions.

* This work was supported by grants from The Danish Cancer Society, the University of Aarhus, Aarhus County, Novo Nordic, the Danish Biotechnology Program, the Frenkels Foundation, the John and Birthe Meyer Foundation, and NCI, National Institutes of Health Grant CA47537. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: Dept. of Clinical Biochemistry, Molecular Diagnostic Laboratory, Aarhus University Hospital, Skejby, DK-8200 Aarhus N, Denmark. Tel.: 45-89495100/45-86156201 (private); Fax: 45-89496018; E-mail: orntoft@kba.sks.au.dk.

REFERENCES

1. Lengauer, C., Kinzler, K. W., and Vogelstein, B. (1998) Genetic instabilities in human cancers. *Nature* **17,** 643–649
2. Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D., and Brown, P. O. (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat. Genet.* **23,** 41–46
3. de Cremoux, P., Martin, E. C., Vincent-Salomon, A., Dieras, V., Barbaroux, C., Liva, S., Pouillart, P., Sastre-Garau, X., and Magdelenat, H. (1999) Quantitative PCR analysis of c-erb B-2 (HER2/neu) gene amplification and comparison with p185(HER2/neu) protein expression in breast cancer drill biopsies. *Int. J. Cancer* **83,** 157–161
4. Brungier, P. P., Tamimi, Y., Shuuring, E., and Schalken, J. (1996) Expression of cyclin D1 and EMS1 in bladder tumors; relationship with chromosome 11q13 amplifications. *Oncogene* **12,** 1747–1753
5. Slavc, I., Ellenbogen, R., Jung, W. H., Vawter, G. F., Kretschmar, C., Grier, H., and Korf, B. R. (1990) *myc* gene amplification and expression in primary human neuroblastoma. *Cancer Res.* **50,** 1459–1463
6. Sauter, G., Carroll, P., Moch, H., Kallioniemi, A., Kerschmann, R., Narayan, P., Mihatsch, M. J., and Waldman, F. M. (1995) c-*myc* copy number gains in bladder cancer detected by fluorescence *in situ* hybridization. *Am J. Pathol.* **146,** 1131–1139
7. Richter, J., Jiang, F., Gorog, J. P., Sartorious, G., Egenter, C., Gasser, T. C., Moch, H., Mihatsch, M. J., and Sauter, G. (1997) Marked genetic differences between stage pTa and stage pT1 papillary bladder cancer detected by comparative genomic hybridization. *Cancer Res.* **57,** 2860–2864
8. Richter, J., Beffa, L., Wagner, U., Schraml, P., Gasser, T. C., Moch, H., Mihatsch, M. J., and Sauter, G. (1998) Patterns of chromosomal imbalances in advanced urinary bladder cancer detected by comparative genomic hybridization. *Am. J. Pathol.* **153,** 1615–1621
9. Bruch, J., Wohr, G., Hautmann, R., Mattfeldt, T., Bruderlein, S., Moller, P., Sauter, S., Hameister, H., Vogel, W., and Paiss, T. (1998) Chromosomal changes during progression of transitional cell carcinoma of the bladder and delineation of the amplified interval on chromosome arm 8q. *Genes Chromosomes Cancer* **23,** 167–174
10. Hovey, R. M., Chu, L., Balazs, M., De Vries, S., Moore, D., Sauter, G., Carroll, P. R., and Waldman, F. M. (1998) Genetic alterations in primary bladder cancers and their metastases. *Cancer Res.* **15,** 3555–3560
11. Simon, R., Burger, H., Brinkschmidt, C., Bocker, W., Hertle, L., and Terpe, H. J. (1998) Chromosomal aberrations associated with invasion in papillary superficial bladder cancer. *J. Pathol.* **185,** 345–351
12. Koo, S. H., Kwon, K. C., Ihm, C. H., Jeon, Y. M., Park, J. W., and Sul, C. K. (1999) Detection of genetic alterations in bladder tumors by comparative genomic hybridization and cytogenetic analysis. *Cancer Genet. Cytogenet.* **110,** 87–93
13. Wodicka, L., Dong, H., Mittmann, M., Ho, M. H., and Lockhart, D. J. (1997) Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **15,** 1359–1367
14. Christensen, M., Sunde, L., Bolund, L., and Orntoft, T. F. (1999) Comparison of three methods of microsatellite detection. *Scand. J. Clin. Lab. Invest.* **59,** 167–177
15. Celis, J. E., Ostergaard, M., Basse, B., Celis, A., Lauridsen, J. B., Ratz, G. P., Andersen, I., Hein, B., Wolf, H., Orntoft, T. F., and Rasmussen, H. H. (1996) Loss of adipocyte-type fatty acid binding protein and other protein biomarkers is associated with progression of human bladder transitional cell carcinomas. *Cancer Res.* **56,** 4782–4790
16. Celis, J. E., Ratz, G., Basse, B., Lauridsen, J. B., and Celis, A. (1994) in *Cell Biology: A Laboratory Handbook* (Celis, J. E., ed) Vol. 3, pp. 222–230, Academic Press, Orlando, FL
17. Ohlsson, R., Tycko, B., and Sapienza, C. (1998) Monoallelic expression: 'there can only be one'. *Trends Genet.* **14,** 435–438
18. Hollander, G. A., Zuklys, S., Morel, C., Mizoguchi, E., Mobisson, K., Simpson, S., Terhorst, C., Wishart, W., Golan, D. E., Bhan, A. K., and Burakoff, S. J. (1998) Monoallelic expression of the interleukin-2 locus. *Science* **279,** 2118–2121
19. Brannan, C. I., and Bartolomei, M. S. (1999) Mechanisms of genomic imprinting. *Curr. Opin. Genet. Dev.* **9,** 164–170
20. Ohlsson, R., Cui, H., He, L., Pfeifer, S., Malmikumpa, H., Jiang, S., Feinberg, A. P., and Hedborg, F. (1999) Mosaic allelic insulin-like growth factor 2 expression patterns reveal a link between Wilms' tumorigenesis and epigenetic heterogeneity. *Cancer Res.* **59,** 3889–3892
21. Cui, H., Hedborg, F., He, L., Nordenskjold, A., Sandstedt, B., Pfeifer-Ohlsson, S., and Ohlsson, R. (1997) Inactivation of H19, an imprinted and putative tumor repressor gene, is a preneoplastic event during Wilms' tumorigenesis. *Cancer Res.* **57,** 4469–4473
22. Galitski, T., Saldanha, A. J., Styles, C. A., Lander, E. S., and Fink, G. R. (1999) Ploidy regulation of gene expression. *Science* **285,** 251–254
23. Tsao, J., Yatabe, Y., Markl, I. D., Hajyan, K., Jones, P. A., and Shibata, D. (2000) Bladder cancer genotype stability during clinical progression. *Genes Chromosomes Cancer* **29,** 26–32
24. Zong, Q., Schummer, M., Hood, L., and Morris, D. R. (1999) Messenger RNA translation state: the second dimension of high-throughput expression screening. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 10632–10636
25. Anderson, L., and Seilhamer, J. (1997) Comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* **18,** 533–537
26. Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., Bumgarner, R., Goodlett, D. R., Aebersold, R., and Hood, L. (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* **292,** 929–934

[CANCER RESEARCH 62, 6240–6245, November 1, 2002]

Impact of DNA Amplification on Gene Expression Patterns in Breast Cancer[1,2]

Elizabeth Hyman,[3] Päivikki Kauraniemi,[3] Sampsa Hautaniemi, Maija Wolf, Spyro Mousses, Ester Rozenblum, Markus Ringnér, Guido Sauter, Outi Monni, Abdel Elkahloun, Olli-P. Kallioniemi, and Anne Kallioniemi[4]

Howard Hughes Medical Institute-NIH Research Scholar, Bethesda, Maryland 20892 [E. H.]; Cancer Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, Maryland 20892 [E. H., P. K., S. H., M. W., S. M., E. R., M. R., A. E., O. K., A. K.]; Laboratory of Cancer Genetics, Institute of Medical Technology, University of Tampere and Tampere University Hospital, FIN-33520 Tampere, Finland [P. K., A. K.]; Signal Processing Laboratory, Tampere University of Technology, FIN-33101 Tampere, Finland [S. H.]; Institute of Pathology, University of Basel, CH-4003 Basel, Switzerland [G. S.]; and Biomedicum Biochip Center, Helsinki University Hospital, Biomedicum Helsinki, FIN-00014 Helsinki, Finland [O. M.]

ABSTRACT

Genetic changes underlie tumor progression and may lead to cancer-specific expression of critical genes. Over 1100 publications have described the use of comparative genomic hybridization (CGH) to analyze the pattern of copy number alterations in cancer, but very few of the genes affected are known. Here, we performed high-resolution CGH analysis on cDNA microarrays in breast cancer and directly compared copy number and mRNA expression levels of 13,824 genes to quantitate the impact of genomic changes on gene expression. We identified and mapped the boundaries of 24 independent amplicons, ranging in size from 0.2 to 12 Mb. Throughout the genome, both high- and low-level copy number changes had a substantial impact on gene expression, with 44% of the highly amplified genes showing overexpression and 10.5% of the highly overexpressed genes being amplified. Statistical analysis with random permutation tests identified 270 genes whose expression levels across 14 samples were systematically attributable to gene amplification. These included most previously described amplified genes in breast cancer and many novel targets for genomic alterations, including the *HOXB7* gene, the presence of which in a novel amplicon at 17q21.3 was validated in 10.2% of primary breast cancers and associated with poor patient prognosis. In conclusion, CGH on cDNA microarrays revealed hundreds of novel genes whose overexpression is attributable to gene amplification. These genes may provide insights to the clonal evolution and progression of breast cancer and highlight promising therapeutic targets.

INTRODUCTION

Gene expression patterns revealed by cDNA microarrays have facilitated classification of cancers into biologically distinct categories, some of which may explain the clinical behavior of the tumors (1–6). Despite this progress in diagnostic classification, the molecular mechanisms underlying gene expression patterns in cancer have remained elusive, and the utility of gene expression profiling in the identification of specific therapeutic targets remains limited.

Accumulation of genetic defects is thought to underlie the clonal evolution of cancer. Identification of the genes that mediate the effects of genetic changes may be important by highlighting transcripts that are actively involved in tumor progression. Such transcripts and their encoded proteins would be ideal targets for anticancer therapies, as demonstrated by the clinical success of new therapies against amplified oncogenes, such as *ERBB2* and *EGFR* (7, 8), in breast cancer and other solid tumors. Besides amplifications of known oncogenes, over 20 recurrent regions of DNA amplification have been mapped in breast cancer by CGH[5] (9, 10). However, these amplicons are often large and poorly defined, and their impact on gene expression remains unknown.

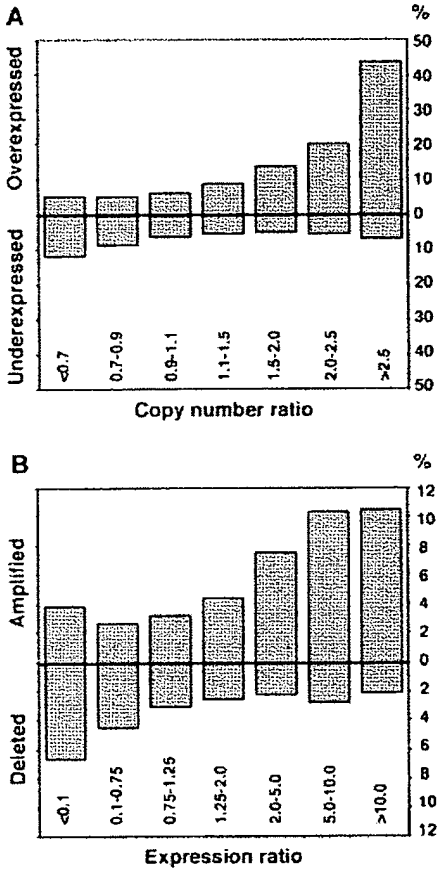


Fig. 1. Impact of gene copy number on global gene expression levels. *A*, percentage of over- and underexpressed genes (*Y axis*) according to copy number ratios (*X axis*). Threshold values used for over- and underexpression were >2.184 (global upper 7% of the cDNA ratios) and <0.4826 (global lower 7% of the expression ratios). *B*, percentage of amplified and deleted genes according to expression ratios. Threshold values for amplification and deletion were >1.5 and <0.7.

We hypothesized that genome-wide identification of those gene expression changes that are attributable to underlying gene copy number alterations would highlight transcripts that are actively involved in the causation or maintenance of the malignant phenotype. To identify such transcripts, we applied a combination of cDNA and CGH microarrays to: (*a*) determine the global impact that gene copy number variation plays in breast cancer development and progression; and (*b*) identify and characterize those genes whose mRNA expres-

Received 5/29/02; accepted 8/28/02.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported in part by the Academy of Finland, Emil Aaltonen Foundation, the Finnish Cancer Society, the Pirkanmaa Cancer Society, the Pirkanmaa Cultural Foundation, the Finnish Breast Cancer Group, the Foundation for the Development of Laboratory Medicine, the Medical Research Fund of the Tampere University Hospital, the Foundation for Commercial and Technical Sciences, and the Swedish Research Council.

[2] Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org).

[3] Contributed equally to this work.

[4] To whom requests for reprints should be addressed, at Laboratory of Cancer Genetics, Institute of Medical Technology, Lenkkeilijankatu 6, FIN-33520 Tampere, Finland. Phone: 358-3247-4125; Fax: 358-3247-4168; E-mail: anne.kallioniemi@uta.fi.

[5] The abbreviations used are: CGH, comparative genomic hybridization; FISH, fluorescence *in situ* hybridization; RT-PCR, reverse transcription-PCR.

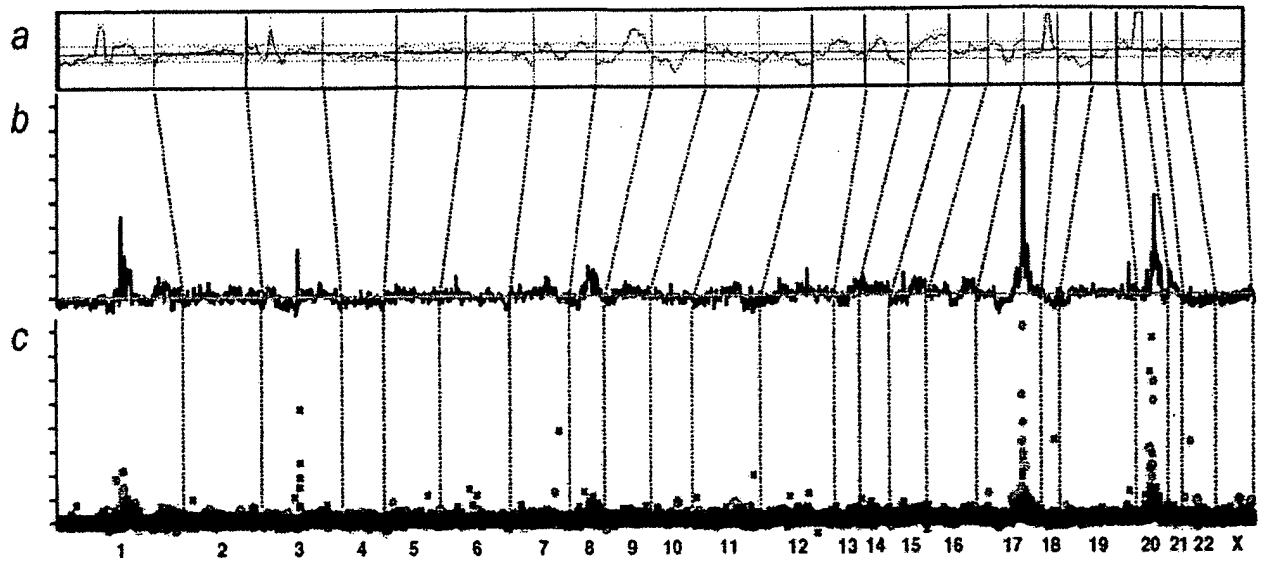


Fig. 2. Genome-wide copy number and expression analysis in the MCF-7 breast cancer cell line. *A*, chromosomal CGH analysis of MCF-7. The copy number ratio profile (*blue line*) across the entire genome from 1p telomere to Xq telomere is shown along with ±1 SD (*orange lines*). The *black horizontal line* indicates a ratio of 1.0; *red line*, a ratio of 0.8; and *green line*, a ratio of 1.2. *B–C*, genome-wide copy number analysis in MCF-7 by CGH on cDNA microarray. The copy number ratios were plotted as a function of the position of the cDNA clones along the human genome. In *B*, individual data points are connected with a line, and a moving median of 10 adjacent clones is shown. *Red horizontal line*, the copy number ratio of 1.0. In *C*, individual data points are labeled by color coding according to cDNA expression ratios. The *bright red dots* indicate the upper 2%, and *dark red dots*, the next 5% of the expression ratios in MCF-7 cells (overexpressed genes); *bright green dots* indicate the lowest 2%, and *dark green dots*, the next 5% of the expression ratios (underexpressed genes); the rest of the observations are shown with *black crosses*. The chromosome numbers are shown at the *bottom* of the figure, and chromosome boundaries are indicated with a *dashed line*.

sion is most significantly associated with amplification of the corresponding genomic template.

MATERIALS AND METHODS

Breast Cancer Cell Lines. Fourteen breast cancer cell lines (BT-20, BT-474, HCC1428, Hs578t, MCF7, MDA-361, MDA-436, MDA-453, MDA-468, SKBR-3, T-47D, UACC812, ZR-75-1, and ZR-75-30) were obtained from the American Type Culture Collection (Manassas, VA). Cells were grown under recommended culture conditions. Genomic DNA and mRNA were isolated using standard protocols.

Copy Number and Expression Analyses by cDNA Microarrays. The preparation and printing of the 13,824 cDNA clones on glass slides were performed as described (11–13). Of these clones, 244 represented uncharacterized expressed sequence tags, and the remainder corresponded to known genes. CGH experiments on cDNA microarrays were done as described (14, 15). Briefly, 20 μg of genomic DNA from breast cancer cell lines and normal human WBCs were digested for 14–18 h with *Alu*I and *Rsa*I (Life Technologies, Inc., Rockville, MD) and purified by phenol/chloroform extraction. Six μg of digested cell line DNAs were labeled with Cy3-dUTP (Amersham Pharmacia) and normal DNA with Cy5-dUTP (Amersham Pharmacia) using the Bioprime Labeling kit (Life Technologies, Inc.). Hybridization (14, 15) and posthybridization washes (13) were done as described. For the expression analyses, a standard reference (Universal Human Reference RNA; Stratagene, La Jolla, CA) was used in all experiments. Forty μg of reference RNA were labeled with Cy3-dUTP and 3.5 μg of test mRNA with Cy5-dUTP, and the labeled cDNAs were hybridized on microarrays as described (13, 15). For both microarray analyses, a laser confocal scanner (Agilent Technologies, Palo Alto, CA) was used to measure the fluorescence intensities at the target locations using the DEARRAY software (16). After background subtraction, average intensities at each clone in the test hybridization were divided by the average intensity of the corresponding clone in the control hybridization. For the copy number analysis, the ratios were normalized on the basis of the distribution of ratios of all targets on the array and for the expression analysis on the basis of 88 housekeeping genes, which were spotted four times onto the array. Low quality measurements (*i.e.*, copy number data with mean reference intensity <100 fluorescent units, and expression data with both test and reference intensity <100 fluorescent units and/or with spot size <50 units) were excluded from the analysis and were treated as missing values. The distributions of fluorescence ratios were used to define cutpoints for increased/decreased copy number. Genes with CGH ratio >1.43 (representing the upper 5% of the CGH ratios across all experiments) were considered to be amplified, and genes with ratio <0.73 (representing the lower 5%) were considered to be deleted.

Statistical Analysis of CGH and cDNA Microarray Data. To evaluate the influence of copy number alterations on gene expression, we applied the following statistical approach. CGH and cDNA calibrated intensity ratios were log-transformed and normalized using median centering of the values in each cell line. Furthermore, cDNA ratios for each gene across all 14 cell lines were median centered. For each gene, the CGH data were represented by a vector that was labeled 1 for amplification (ratio, >1.43) and 0 for no amplification. Amplification was correlated with gene expression using the signal-to-noise statistics (1). We calculated a weight, w_g, for each gene as follows:

$$w_g = \frac{m_{g1} - m_{g0}}{\sigma_{g1} + \sigma_{g0}}$$

where m_{g1}, σ_{g1} and m_{g0}, σ_{g0} denote the means and SDs for the expression levels for amplified and nonamplified cell lines, respectively. To assess the statistical significance of each weight, we performed 10,000 random permutations of the label vector. The probability that a gene had a larger or equal weight by random permutation than the original weight was denoted by α. A low α (<0.05) indicates a strong association between gene expression and amplification.

Genomic Localization of cDNA Clones and Amplicon Mapping. Each cDNA clone on the microarray was assigned to a Unigene cluster using the Unigene Build 141.[6] A database of genomic sequence alignment information for mRNA sequences was created from the August 2001 freeze of the University of California Santa Cruz's GoldenPath database.[7] The chromosome and bp positions for each cDNA clone were then retrieved by relating these data sets. Amplicons were defined as a CGH copy number ratio >2.0 in at least two adjacent clones in two or more cell lines or a CGH ratio >2.0 in at least three adjacent clones in a single cell line. The amplicon start and end positions were

[6] Internet address: http://research.nhgri.nih.gov/microarray/downloadable_cdna.html.
[7] Internet address: www.genome.ucsc.edu.

Table 1 *Summary of independent amplicons in 14 breast cancer cell lines by CGH microarray*

Location	Start (Mb)	End (Mb)	Size (Mb)
1p13	132.79	132.94	0.2
1q21	173.92	177.25	3.3
1q22	179.28	179.57	0.3
3p14	71.94	74.66	2.7
7p12.1–7p11.2	55.62	60.95	5.3
7q31	125.73	130.96	5.2
7q32	140.01	140.68	0.7
8q21.11–8q21.13	86.45	92.46	6.0
8q21.3	98.45	103.05	4.6
8q23.3–8q24.14	129.88	142.15	12.3
8q24.22	151.21	152.16	1.0
9p13	38.65	39.25	0.6
13q22–q31	77.15	81.38	4.2
16q22	86.70	87.62	0.9
17q11	29.30	30.85	1.6
17q12–q21.2	39.79	42.80	3.0
17q21.32–q21.33	52.47	55.80	3.3
17q22–q23.3	63.81	69.70	5.9
17q23.3–q24.3	69.93	74.99	5.1
19q13	40.63	41.40	0.8
20q11.22	34.59	35.85	1.3
20q13.12	44.00	45.62	1.6
20q13.12–q13.13	46.45	49.43	3.0
20q13.2–q13.32	51.32	59.12	7.8

extended to include neighboring nonamplified clones (ratio, <1.5). The amplicon size determination was partially dependent on local clone density.

FISH. Dual-color interphase FISH to breast cancer cell lines was done as described (17). Bacterial artificial chromosome clone RP11-361K8 was labeled with SpectrumOrange (Vysis, Downers Grove, IL), and SpectrumOrange-labeled probe for *EGFR* was obtained from Vysis. SpectrumGreen-labeled chromosome 7 and 17 centromere probes (Vysis) were used as a reference. A tissue microarray containing 612 formalin-fixed, paraffin-embedded primary breast cancers (17) was applied in FISH analyses as described (18). The use of these specimens was approved by the Ethics Committee of the University of Basel and by the NIH. Specimens containing a 2-fold or higher increase in the number of test probe signals, as compared with corresponding centromere signals, in at least 10% of the tumor cells were considered to be amplified. Survival analysis was performed using the Kaplan-Meier method and the log-rank test.

RT-PCR. The *HOXB7* expression level was determined relative to *GAPDH*. Reverse transcription and PCR amplification were performed using Access RT-PCR System (Promega Corp., Madison, WI) with 10 ng of mRNA as a template. *HOXB7* primers were 5′-GAGCAGAGGGACTCGGACTT-3′ and 5′-GCGTCAGGTAGCGATTGTAG-3′.

RESULTS

Global Effect of Copy Number on Gene Expression. 13,824 arrayed cDNA clones were applied for analysis of gene expression and gene copy number (CGH microarrays) in 14 breast cancer cell lines. The results illustrate a considerable influence of copy number on gene expression patterns. Up to 44% of the highly amplified transcripts (CGH ratio, >2.5) were overexpressed (*i.e.*, belonged to the global upper 7% of expression ratios), compared with only 6% for genes with normal copy number levels (Fig. 1A). Conversely, 10.5% of the transcripts with high-level expression (cDNA ratio, >10) showed increased copy number (Fig. 1*B*). Low-level copy number increases and decreases were also associated with similar, although less dramatic, outcomes on gene expression (Fig. 1).

Identification of Distinct Breast Cancer Amplicons. Base-pair locations obtained for 11,994 cDNAs (86.8%) were used to plot copy number changes as a function of genomic position (Fig. 2, Supplement Fig. A). The average spacing of clones throughout the genome was 267 kb. This high-resolution mapping identified 24 independent breast cancer amplicons, spanning from 0.2 to 12 Mb of DNA (Table 1). Several amplification sites detected previously by chromosomal CGH were validated, with 1q21, 17q12–q21.2, 17q22–q23, 20q13.1, and 20q13.2 regions being most commonly amplified. Furthermore, the boundaries of these amplicons were precisely delineated. In addition, novel amplicons were identified at 9p13 (38.65–39.25 Mb), and 17q21.3 (52.47–55.80 Mb).

Direct Identification of Putative Amplification Target Genes. The cDNA/CGH microarray technique enables the direct correlation of copy number and expression data on a gene-by-gene basis throughout the genome. We directly annotated high-resolution CGH plots with gene expression data using color coding. Fig. 2*C* shows that most of the amplified genes in the MCF-7 breast cancer cell line at 1p13, 17q22–q23, and 20q13 were highly overexpressed. A view of chromosome 7 in the MDA-468 cell line implicates *EGFR* as the most highly overexpressed and amplified gene at 7p11–p12 (Fig. 3*A*). In BT-474, the two known amplicons at 17q12 and 17q22–q23 contained numerous highly overexpressed genes (Fig. 3*B*). In addition, several genes, including the homeobox genes *HOXB2* and *HOXB7*, were highly amplified in a previously undescribed independent amplicon at 17q21.3. *HOXB7* was systematically amplified (as validated by FISH, Fig. 3*B*, *inset*) as well as overexpressed (as verified by RT-PCR, data not shown) in BT-474, UACC812, and ZR-75-30 cells. Furthermore, this novel

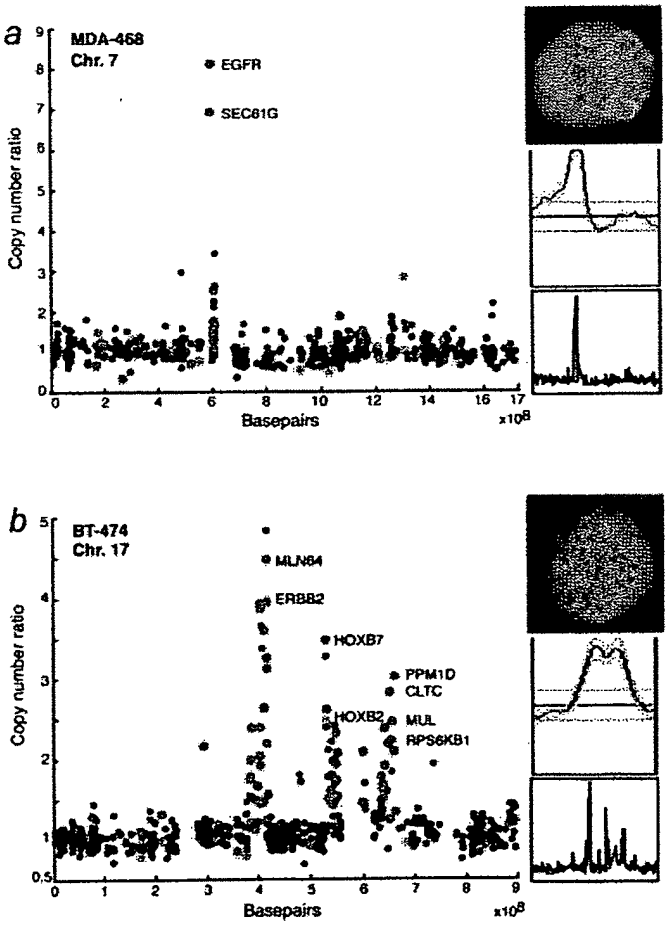


Fig. 3. Annotation of gene expression data on CGH microarray profiles. *A*, genes in the 7p11-p12 amplicon in the MDA-468 cell line are highly expressed (*red dots*) and include the *EGFR* oncogene. *B*, several genes in the 17q12, 17q21.3, and 17q23 amplicons in the BT-474 breast cancer cell line are highly overexpressed (*red*) and include the *HOXB7* gene. The data labels and color coding are as indicated for Fig. 2*C*. *Insets* show chromosomal CGH profiles for the corresponding chromosomes and validation of the increased copy number by interphase FISH using *EGFR* (*red*) and chromosome 7 centromere probe (*green*) to MDA-468 (*A*) and *HOXB7*-specific probe (*red*) and chromosome 17 centromere (*green*) to BT-474 cells (*B*).

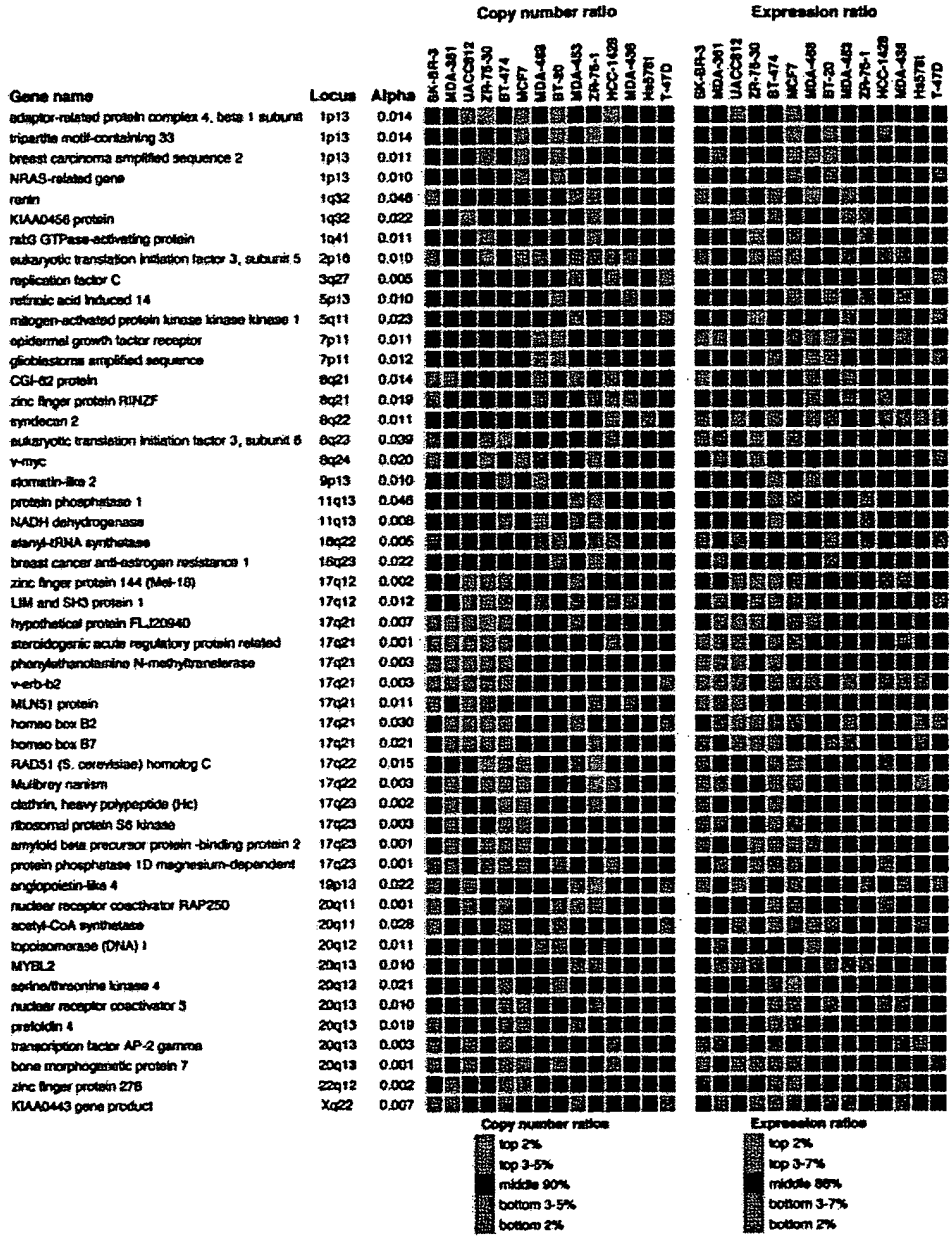


Fig. 4. List of 50 genes with a statistically significant correlation (α value <0.05) between gene copy number and gene expression. Name, chromosomal location, and the α value for each gene are indicated. The genes have been ordered according to their position in the genome. The color maps on the *right* illustrate the copy number and expression ratio patterns in the 14 cell lines. The key to the color code is shown at the *bottom* of the graph. *Gray squares*, missing values. The complete list of 270 genes is shown in supplemental Fig. B.

amplification was validated to be present in 10.2% of 363 primary breast cancers by FISH to a tissue microarray and was associated with poor prognosis of the patients ($P = 0.001$).

Statistical Identification and Characterization of 270 Highly Expressed Genes in Amplicons. Statistical comparison of expression levels of all genes as a function of gene amplification identified 270 genes whose expression was significantly influenced by copy number across all 14 cell lines (Fig. 4, Supplemental Fig. B). According to the gene ontology data,[8] 91 of the 270 genes represented hypothetical proteins or genes with no functional annotation, whereas 179 had associated functional information available. Of these, 151 (84%) are implicated in apoptosis, cell proliferation, signal transduction, and transcription, whereas 28 (16%) had functional annotations that could not be directly linked with cancer.

DISCUSSION

The importance of recurrent gene and chromosome copy number changes in the development and progression of solid tumors has been characterized in >1000 publications applying CGH[9] (9, 10), as well as in a large number of other molecular cytogenetic, cytogenetic, and molecular genetic studies. The effects of these somatic genetic changes on gene expression levels have remained largely unknown, although a few studies have explored gene expression changes occurring in specific amplicons (15, 19–21). Here, we applied genome-wide cDNA microarrays to identify transcripts whose expression changes were attributable to underlying gene copy number alterations in breast cancer.

The overall impact of copy number on gene expression patterns was substantial with the most dramatic effects seen in the case of high-

[8] Internet address: http://www.geneontology.org/.

[9] Internet address: http://www.ncbi.nlm.nih.gov/entrez.

level copy number increase. Low-level copy number gains and losses also had a significant influence on expression levels of genes in the regions affected, but these effects were more subtle on a gene-by-gene basis than those of high-level amplifications. However, the impact of low-level gains on the dysregulation of gene expression patterns in cancer may be equally important if not more important than that of high-level amplifications. Aneuploidy and low-level gains and losses of chromosomal arms represent the most common types of genetic alterations in breast and other cancers and, therefore, have an influence on many genes. Our results in breast cancer extend the recent studies on the impact of aneuploidy on global gene expression patterns in yeast cells, acute myeloid leukemia, and a prostate cancer model system (22–24).

The CGH microarray analysis identified 24 independent breast cancer amplicons. We defined the precise boundaries for many amplicons detected previously by chromosomal CGH (9, 10, 25, 26) and also discovered novel amplicons that had not been detected previously, presumably because of their small size (only 1–2 Mb) or close proximity to other larger amplicons. One of these novel amplicons involved the homeobox gene region at 17q21.3 and led to the overexpression of the *HOXB7* and *HOXB2* genes. The homeodomain transcription factors are known to be key regulators of embryonic development and have been occasionally reported to undergo aberrant expression in cancer (27, 28). *HOXB7* transfection induced cell proliferation in melanoma, breast, and ovarian cancer cells and increased tumorigenicity and angiogenesis in breast cancer (29–32). The present results imply that gene amplification may be a prominent mechanism for overexpressing *HOXB7* in breast cancer and suggest that *HOXB7* contributes to tumor progression and confers an aggressive disease phenotype in breast cancer. This view is supported by our finding of amplification of *HOXB7* in 10% of 363 primary breast cancers, as well as an association of amplification with poor prognosis of the patients.

We carried out a systematic search to identify genes whose expression levels across all 14 cell lines were attributable to amplification status. Statistical analysis revealed 270 such genes (representing ~2% of all genes on the array), including not only previously described amplified genes, such as *HER-2*, *MYC*, *EGFR*, ribosomal protein s6 kinase, and *AIB3*, but also numerous novel genes such as *NRAS-related gene* (1p13), *syndecan-2* (8q22), and *bone morphogenic protein* (20q13.1), whose activation by amplification may similarly promote breast cancer progression. Most of the 270 genes have not been implicated previously in breast cancer development and suggest novel pathogenetic mechanisms. Although we would not expect all of them to be causally involved, it is intriguing that 84% of the genes with associated functional information were implicated in apoptosis, cell proliferation, signal transduction, transcription, or other cellular processes that could directly imply a possible role in cancer progression. Therefore, a detailed characterization of these genes may provide biological insights to breast cancer progression and might lead to the development of novel therapeutic strategies.

In summary, we demonstrate application of cDNA microarrays to the analysis of both copy number and expression levels of over 12,000 transcripts throughout the breast cancer genome, roughly once every 267 kb. This analysis provided: (*a*) evidence of a prominent global influence of copy number changes on gene expression levels; (*b*) a high-resolution map of 24 independent amplicons in breast cancer; and (*c*) identification of a set of 270 genes, the overexpression of which was statistically attributable to gene amplification. Characterization of a novel amplicon at 17q21.3 implicated amplification and overexpression of the *HOXB7* gene in breast cancer, including a clinical association between *HOXB7* amplification and poor patient prognosis. Overall, our results illustrate how the identification of genes activated by gene amplification provides a powerful approach to highlight genes with an important role in cancer as well as to prioritize and validate putative targets for therapy development.

REFERENCES

1. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science (Wash. DC), *286:* 531–537, 1999.
2. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature (Lond.), *403:* 503–511, 2000.
3. Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben-Dor, A., *et al.* Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature (Lond.), *406:* 536–540, 2000.
4. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., *et al.* Molecular portraits of human breast tumours. Nature (Lond.), *406:* 747–752, 2000.
5. Dhanasekaran, S. M., Barrette, T. R., Ghosh, D., Shah, R., Varambally, S., Kurachi, K., Pienta, K. J., Rubin, M. A., and Chinnaiyan, A. M. Delineation of prognostic biomarkers in prostate cancer. Nature (Lond.), *412:* 822–826, 2001.
6. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc. Natl. Acad. Sci. USA, *98:* 10869–10874, 2001.
7. Ross, J. S., and Fletcher, J. A. The *HER-2/neu* oncogene: prognostic factor, predictive factor and target for therapy. Semin. Cancer Biol., *9:* 125–138, 1999.
8. Arteaga, C. L. The epidermal growth factor receptor: from mutant oncogene in nonhuman cancers to therapeutic target in human neoplasia. J. Clin. Oncol., *19:* 32–40, 2001.
9. Knuutila, S., Bjorkqvist, A. M., Autio, K., Tarkkanen, M., Wolf, M., Monni, O., Szymanska, J., Larramendy, M. L., Tapper, J., Pere, H., El-Rifai, W., *et al.* DNA copy number amplifications in human neoplasms: review of comparative genomic hybridization studies. Am. J. Pathol., *152:* 1107–1123, 1998.
10. Knuutila S., Autio K., and Aalto Y. Online access to CGH data of DNA sequence copy number changes. Am. J. Pathol., *157:* 689, 2000.
11. DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A., and Trent, J. M. Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat. Genet., *14:* 457–460, 1996.
12. Shalon, D., Smith, S. J., and Brown, P. O. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. Genome Res., *6:* 639–645, 1996.
13. Mousses, S., Bittner, M. L., Chen, Y., Dougherty, E. R., Baxevanis, A., Meltzer, P. S., and Trent, J. M. Gene expression analysis by cDNA microarrays. *In:* F. J. Livesey and S. P. Hunt (eds.), Functional Genomics, pp. 113–137. Oxford: Oxford University Press, 2000.
14. Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D., and Brown, P. O. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat. Genet., *23:* 41–46, 1999.
15. Monni, O., Bärlund, M., Mousses, S., Kononen, J., Sauter, G., Heiskanen, M., Paavola, P., Avela, K., Chen, Y., Bittner, M. L., and Kallioniemi, A. Comprehensive copy number and gene expression profiling of the 17q23 amplicon in human breast cancer. Proc. Natl. Acad. Sci. USA, *98:* 5711–5716, 2001.
16. Chen, Y., Dougherty, E. R., and Bittner, M. L. Ratio-based decisions and the quantitative analysis of cDNA microarray images. J. Biomed. Optics, *2:* 364–374, 1997.
17. Bärlund, M., Forozan, F., Kononen, J., Bubendorf, L., Chen, Y., Bittner, M. L., Torhorst, J., Haas, P., Bucher, C., Sauter, G., *et al.* Detecting activation of ribosomal protein S6 kinase by complementary DNA and tissue microarray analysis. J. Natl. Cancer Inst., *92:* 1252–1259, 2000.
18. Andersen, C. L., Hostetter, G., Grigoryan, A., Sauter, G., and Kallioniemi, A. Improved procedure for fluorescence *in situ* hybridization on tissue microarrays. Cytometry, *45:* 83–86, 2001.
19. Kauraniemi, P., Bärlund, M., Monni, O., and Kallioniemi, A. New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNA microarrays. Cancer Res., *61:* 8235–8240, 2001.
20. Clark, J., Edwards, S., John, M., Flohr, P., Gordon, T., Maillard, K., Giddings, I., Brown, C., Bagherzadeh, A., Campbell, C., Shipley, J., Wooster, R., and Cooper, C. S. Identification of amplified and expressed genes in breast cancer by comparative hybridization onto microarrays of randomly selected cDNA clones. Genes Chromosomes Cancer, *34:* 104–114, 2002.
21. Varis, A., Wolf, M., Monni, O., Vakkari, M. L., Kokkola, A., Moskaluk, C., Frierson, H., Powell, S. M., Knuutila, S., Kallioniemi, A., and El-Rifai, W. Targets of gene amplification and overexpression at 17q in gastric cancer. Cancer Res., *62:* 2625–2629, 2002.
22. Hughes, T. R., Roberts, C. J., Dai, H., Jones, A. R., Meyer, M. R., Slade, D., Burchard, J., Dow, S., Ward, T. R., Kidd, M. J., Friend, S. H., and Marton M. J.

Widespread aneuploidy revealed by DNA microarray expression profiling. Nat. Genet., *25:* 333–337, 2000.

23. Virtaneva, K., Wright, F. A., Tanner, S. M., Yuan, B., Lemon, W. J., Caligiuri, M. A., Bloomfield, C. D., de La Chapelle, A., and Krahe, R. Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics. Proc. Natl. Acad. Sci. USA, *98:* 1124–1129, 2001.
24. Phillips, J. L., Hayward, S. W., Wang, Y., Vasselli, J., Pavlovich, C., Padilla-Nash, H., Pezullo, J. R., Ghadimi, B. M., Grossfeld, G. D., Rivera, A., Linehan, W. M., Cunha, G. R., and Ried, T. The consequences of chromosomal aneuploidy on gene expression profiles in a cell line model for prostate carcinogenesis. Cancer Res., *61:* 8143–8149, 2001.
25. Bärlund, M., Tirkkonen, M., Forozan, F., Tanner, M. M., Kallioniemi, O. P., and Kallioniemi, A. Increased copy number at 17q22–q24 by CGH in breast cancer is due to high-level amplification of two separate regions. Genes Chromosomes Cancer, *20:* 372–376, 1997.
26. Tanner, M. M., Tirkkonen, M., Kallioniemi, A., Isola, J., Kuukasjärvi, T., Collins, C., Kowbel, D., Guan, X. Y., Trent, J., Gray, J. W., Meltzer, P., and Kallioniemi O. P. Independent amplification and frequent co-amplification of three nonsyntenic regions on the long arm of chromosome 20 in human breast cancer. Cancer Res., *56:* 3441–3445, 1996.
27. Cillo, C., Faicila, A., Cantile, M., and Boncinelli, E. Homeobox genes and cancer. Exp. Cell Res., *248:* 1–9, 1999.
28. Cillo, C., Cantile, M., Faicila, A., and Boncinelli, E. Homeobox genes in normal and malignant cells. J. Cell. Physiol., *188:* 161–169, 2001.
29. Care, A., Silvani, A., Meccia, E., Mattia, G., Stoppacciaro, A., Parmiani, G., Peschle, C., and Colombo, M. P. HOXB7 constitutively activates basic fibroblast growth factor in melanomas. Mol. Cell. Biol., *16:* 4842–4851, 1996.
30. Care, A., Silvani, A., Meccia, E., Mattia, G., Peschle, C., and Colombo, M. P. Transduction of the SkBr3 breast carcinoma cell line with the *HOXB7* gene induces bFGF expression, increases cell proliferation and reduces growth factor dependence. Oncogene, *16:* 3285–3289, 1998.
31. Care, A., Felicetti, F., Meccia, E., Bottero, L., Parenza, M., Stoppacciaro, A., Peschle, C., and Colombo, M. P. HOXB7: a key factor for tumor-associated angiogenic switch. Cancer Res., *61:* 6532–6539, 2001.
32. Naora, H., Yang, Y. Q., Montz, F. J., Seidman, J. D., Kurman, R. J., and Roden, R. B. A serologically identified tumor antigen encoded by a homeobox gene promotes growth of ovarian epithelial cells. Proc. Natl. Acad. Sci. USA, *98:* 4060–4065, 2001.

Breast Cancer Research and Treatment **63**: 163–169, 2000.

Report

Relevance of p185 HER-2/neu oncoprotein quantification in human primary breast carcinoma

Laurent Bermont[1], Marie-Paule Algros[2], Marie-Hélène Baron[3], and Gérard L. Adessi[1]
[1]*Service d'Oncologie et d'Endocrinologie Moléculaires (C.H.U. Besançon) and C.R.I 96.01, Besançon, France;* [2]*Service d'Anatomie et Cytologie Pathologique, C.H.U. Besançon, France;* [3]*Service de Radiothérapie et Oncologie, C.H.U. Besançon, France*

Key words: breast carcinoma, *c-erbB-2* amplification, p185 overexpression, prognostic factors

Summary

The *c-erbB-2* proto-oncogene encodes a transmembrane protein tyrosine kinase receptor of 185 kDa (p185) and has been associated with several types of human cancers. In human breast cancer, overexpression of p185 occurs in 15–30% of cases, correlates with poor prognostic factors and characterizes breast cancers with a more aggressive behavior. Overexpression of p185 is usually associated with *c-erbB-2* amplification, though it may occur independently and thus define subpopulations of breast cancers which might be of clinical interest. p185 expression is usually detected by immunohistochemistry (IHC) and few studies have been carried out to evaluate the p185 content of breast cancers with an ELISA technique. In this context, we showed, in 106 breast cancer samples, that p185 was expressed at high levels in 13.2%, intermediate levels in 55.7% and negative ones in 31.1% of cases. All p185 positive samples showed a *c-erbB-2* oncogene amplification while none of the p185 negative samples and only 4% of p185 imtermediate samples had an amplification of *c-erbB-2*. p185 expression is significantly correlated with the negativity of estrogen and progestrone receptors, with high levels of cathepsin D and in some conditions with axillary nodal involvement. Thus, using the p185 ELISA assay, the *c-erbB-2* status of breast cancers can be defined and moreover a subset can be discriminated which is characterized by intermediate levels of p185 and absence of *c-erbB-2* amplification. The quantitative approach towards p185 in breast cancers affords the possibility of identifying more appropriately patients with high or low risk and thus permits adaptation of therapeutic regimens.

Introduction

The *c-erbB-2* proto-oncogene encodes a transmembrane protein tyrosine kinase receptor of 185 kDa (p185) having extensive homology with the epidermal growth factor (EGF) receptor, also known as c-erbB-1 [1]. Both these oncogenes belong to a family, also including c-erbB-3 and c-erbB-4 [2], whose alterations have been associated with several types of human cancers, particularly those involving the breast [3]. The neu oncogene (which is the rat homologue of *c-erbB-2*) was initially identified in chemically induced rat neuroblastoma as the active transforming element [4, 5] and oncogenic activation of Neu occurs as the result of a single point mutation in the transmembrane domain ($Val^{664} \rightarrow Glu$) [6]. Overexpression of p185, as well as specific point mutations in the transmembrane domain, cause an increase of the tyrosine kinase activity of Neu and have been shown to have a transforming effect in cells [7, 8].

Although transmembrane domain point mutations are apparently rare in human tumor, overexpression of p185 has been identified in a wide range of human cancers and especially in breast and ovarian tumors [9]. In human breast cancer, overexpression of p185 occurs in 15–30% of cases and predicts a significantly lower survival rate and a shorter time before relapse in patients with lymph node-positive disease [9, 10]. Overexpression of p185 is usually associated with *c-erbB-2* amplification, though it may occur in-

dependently and thus define subpopulations of breast cancers which might be of clinical interest. This fact shows that whatever the *c-erbB-2* gene status, overexpression of p185 is the result of a transcriptional and/or translational dysregulation and evaluation of the p185 oncoprotein appears to be more appropriate for characterizing tumoral subpopulations than *c-erbB-2* amplification.

In previous studies, we reported an amplification of the *c-erbB-2* oncogene in 15–20% of human breast cancer samples [11, 12] which is in agreement with the 10–34% of breast cancers having a *c-erbB-2* amplification [13, 14]. However, the evaluation of p185 expression might constitute a better approach, than *c-erbB-2* gene amplification, regarding the predictive value of this parameter in breast cancer evolution. Thus, the aim of our study was to define subsets of breast cancers in terms of p185 expression and to determine the relationship of p185 expression with the accepted prognostic factors in a breast cancer population.

Materials and methods

Patients

Primary breast carcinoma samples were obtained from 106 women undergoing surgery, from July 1995 to December 1995, at the Medical School Hospital (Besançon, France). For each patient, several clinical and pathological parameters were collected afterwards: age, tumor size, the Scarf Bloom Richardson histological grading (SBR grading), axillary nodal involvement, hormonal receptors and cathepsin D.

DNA preparation

The tumor samples were frozen and kept in liquid nitrogen until processing. The tumor DNA was collected from the remaining cellular pellet after separation of the cytosol fraction used for the hormonal receptor assay. DNA obtained from the leukocytes of healthy donors was used as reference DNA. Total genomic DNA extraction was carried out using the Genomix method (Kontron Instrument, Milan, Italy) according to the manufacturer's instructions.

Slot blot hybridization analysis

Analysis of *c-erbB-2* amplification was performed as previously described [11, 12]. For each sample, three DNA blots (10, 5 and 2.5 μg) were performed and hybridized with a *c-erbB-2* probe corresponding to a 1.03 kb *Eco*RI fragment of pGneu1 (Dr O Brison, Institut Gustave Roussy, Villejuif, France). Filters were subsequently hybridized with a β-globin probe which was a 1.9 kb *Bam*HI fragment of pSP64 (Pr M Goosens, Inserm U91, Paris, France) for normalization. Both probes were labelled with digoxigenin-11-dUTP by Nick-Translation according to the recommendations of the manufacturer (Boehringer Mannheim). After washing, labelled probes were revealed by a chemiluminescent reaction (DIG Luminescent Detection Kit, Boehringer Mannheim) as described by the manufacturer. Then filters were exposed for autoradiography and signals were quantified using a laser densitometer Ultroscan XL™ (Pharmacia LKB). The degree of c-erbB-2 amplification was evaluated by the ratio of the corrected signal for the tumor DNA to the corrected signal for the leukocyte DNA. A tumor was considered to be amplified when the corrected ratio was at least twice that of the leukocyte DNA.

Immunoanalysis

Human p185 oncoprotein was measured by enzyme immunoassay using the Human Neu Quantitative ELISA Assay™ kit (Oncogene Science) on an aliquot of cell lysates before separation of the cytosol fraction used for the hormonal receptor assay. The p185 protein ELISA assay was performed using two antibodies (NB3 and TA1) directed against the extracellular domain of the protein. Cuny et al. (1994), using the same immunoassay, have shown that the maximum value of p185 in non-tumoral breast tissue samples was 3 U/μg protein (1 U of p185 = 0.05 fmol) and this negative cutoff value was retained to define the p185 negative subgroup. A positive cutoff value does not seem to have been definitively established and differs according to the manufacturers. A positive cutoff value of 20 U of p185 per microgram protein, which defines the p185 positive subgroup, was retained as the most informative positive cutoff value with respect to the correlation with the clinical and biological parameters. Both these cutoff values define a third subgroup of intermediate p185 (between 3 and 20 U/μg protein).

Cytosolic estradiol receptor (ER) and progesterone receptor (PgR) determinations were performed by enzyme immunoassay ER-EIA and PgR-EIA kits (Abbott, France). As established by the manufacturer, all values above 15 fmol/mg protein were considered to be receptor positive.

Cytosolic cathepsin D was measured by immunoradiometric assay ELSA-Cath-D (Cis Biointernational, France). A level of cathepsin D greater than 60 pmol/mg protein was considered to be pathologic, and a level below 30 pmol/mg protein was considered to be normal. Both these cutoff values define a third subgroup between 30 and 60 pmol/mg protein where the prognostic value of cathepsin D is more difficult to establish [15].

Statistical analysis

A univariate analysis (χ^2-test) was used to compare the distribution of the main parameters describing the patients and the tumors.

A Student's test was used to compare the mean concentrations. The difference was considered significant when the p value was less than 0.05.

Results

The lack of difference between the population studied and a population of 550 breast cancers, explored during the same period of time, demonstrates the absence of either an intentional or an unintentional selection bias and shows that this population can be considered as representative of the general population (Table 1).

p185 expression correlates with c-erbB-2 amplification

Immunoenzymatic assay of p185, performed in 106 breast cancer samples, showed a p185 which was positive in 14 cases (13.2%), intermediate in 59 cases (55.7%) and negative in 33 cases (31.1%). Slot blot analysis for *c-erbB-2* amplification was performed in 89 of these 106 samples and a *c-erbB-2* amplification was detected in 14 samples (15.7%), ranging from two to 20 copies. The distribution of *c-erbB-2* amplification in terms of p185 expression is shown in Table 2 and there is a positive correlation between the degree of *c-erbB-2* amplification and the level of p185 expression ($R^2 = 0.85$). The p185 intermediate subgroup (52.8% of cases) is characterized by absence of gene amplification in 47/49 cases (~96%) and there is no amplification when p185 is negative (<3 U/μg protein). Thus, the cutoff values of p185 are suitable for discriminating breast cancers having a *c-erbB-2* amplification from those without amplification.

Table 1. Population comparison between a reference population and the population studied

Parameters		Reference population ($n = 550$)	p185 population ($n = 106$)	p^1
Age at diagnosis	<50	135	31	0.31
(years)	>50	415	75	
Estrogen	<15	159	27	0.94
receptors	>15	391	79	
(fmol/mg prot.)				
Progesterone	<15	180	32	0.63
receptors	>15	370	74	
(fmol/mg prot.)				
Cathepsin D	<30	125	18	0.26
(pmol/mg prot.)	30 to 60	172	40	
	>60	203	45	
SBR histological	I	73	7	0.16
grade[2]	II	301	57	
	III	126	28	
Axillary hodal	N_0	213	34	0.36
involvement[3]	N+	157	32	

[1] χ^2-test, the difference was considered significant when the p value was less than 0.05.
[2] SBR grading according to Scarff, Bloom and Richardson method.
[3] N_0 sub-group encompass patients without axillary nodal involvement while N+ sub-group those with axillary nodal involvement.

Table 2. Distribution of *c-erbB-2* oncogene amplification and p185 oncoprotein expression in a population of breast carcinomas

		p185 expression		
		Positive	Intermediate	Negative
c-erbB-2	+	12	2	0
amplification	–	0	47	28

Relationship between p185 expression and clinical and biological parameters

Table 3 summarizes the correlation results and shows that p185 expression correlates with the estrogen receptors ($p = 0.023$), the progesterone receptors ($p = 0.036$) and the cathepsin D ($p = 0.027$). In order to better define the correlation of p185 expression with the steroid receptors, we determined the mean concentrations of both steroid receptors for each p185 subgroup (Table 4). The results show that the p185 intermediate subgroup expresses significantly higher levels of both steroid receptors as compared to the p185 positive subgroup. Surprisingly, the p185 negative subgroup also expresses low levels of estrogen

Table 3. Relationship between p185 expression and clinical and biological parameters in a population of breast carcinomas

Parameters		p185 positive ($n = 14$)	p185 intermediate ($n = 59$)	p185 negative ($n = 33$)	p^1
Age at diagnosis	<50	7	16	8	0.179
(years)	>50	7	43	25	
Estrogen	<15	6	9	12	0.023
receptors	>15	8	50	21	
(fmol/mg prot.)					
Progesterone	<15	7	12	13	0.036
receptors	>15	7	47	20	
Cathepsin D	<30	0	7	11	0.027
(pmol/mg prot.)	30 to 60	4	26	10	
	>60	8	25	12	
Histological	Ductal	14	56	27	0.152
type	lobular	0	3	6	
SBR histological	I	2	4	1	0.494
grade[2]	II	6	38	13	
	III	6	12	10	
Axillary nodal	N_0	3	16	15	0.289[4]
involvement[3]	N+	5	20	7	
Size (cm)	<1	2	6	6	0.620
	>1	12	52	24	

[1] χ^2-test, the difference was considered significant when the p value was less than 0.05.
[2] SBR grading according to Scarff, Bloom and Richardson method.
[3] N_0 sub-group encompass patients without axillary nodal involvement while N+ sub-group those with axillary nodal involvement.
[4] when p185 positive and intermediate are pooled and compared with p185 negative then $p = 0.055$.

Table 4. Relationship between p185 expression and steroid receptors in a population of breast carcinomas

Parameters		n	Estrogen receptors (fmol/mg prot.)	p^1	Progesterone receptors (fmol/mg prot.)	p^1
p185 expression	<3	33	72 ± 27	0.000007[2]	224 ± 128	0.073[2]
(U/μg prot.)	3–20	59	249 ± 65		403 ± 141	
	>20	14	61 ± 68	0.00047[3]	127 ± 103	0.0036[3]

[1] Student's test, the difference was considered significant when the p value was less than 0.05.
[2] p185 <3 vs. p185 3–20.
[3] p185 3–20 vs. p185 >20, (results are expressed as the mean ± sem).

receptors which are not significantly different from those observed in the p185 positive subgroup.

For the other parameters, there is no correlation with p185 expression when the two cutoff values system is applied. On the other hand, a positive cutoff value of 3 U of p185 per microgram protein improves the correlation with the axillary nodal involvement ($p = 0.055$). In this case statistical significance was reached when a positive cutoff value of 4 U of p185 per microgram protein ($p = 0.025$) was applied and confirms the correlation of p185 overexpression with the axillary nodal involvement. However, there is a decrease in the statistical significance of correlations between p185 expression and the other clinical and

biological parameters when a positive cutoff value of 4 U of p185 per microgram protein is used.

Thus, according to the positive cutoff value of p185, the correlation between the p185 expression and the clinical and biological parameters fluctuates. These fluctuations result from the ambiguous behavior of the p185 intermediate subgroup which is sometimes associated with good prognostic factors (estrogen and progesterone receptors) like the p185 negative subgroup, and is sometimes associated with poor prognostic factors (cathepsin D and axillary nodal involvement) like the p185 positive subgroup.

In our breast cancer population, no correlation of p185 expression with the histological type or grade was observed. In addition, breast carcinomas of lobular type never overexpressed p185, and only 3/9 (33.3%) expressed p185 at intermediate levels (Table 3).

Discussion

c-erbB-2 proto-oncogene amplification occurs in 10–35% of breast cancers and constitutes a poor prognostic factor [9, 10]. The main expression product of this gene is a 185 kDa transmembrane protein whose level of expression depends on (i) the degree of amplification of the gene and/or, (ii) an increase of the transcription rate of the gene. Therefore, a substantial proportion of breast cancers express high or moderate levels of p185 without *c-erbB-2* oncogene amplification. Characterization of this subset of breast cancers is of clinical interest and might thus be helpful for clinicians in the treatment decision process.

In breast cancers, numerous studies have been performed to investigate the *c-erbB-2* amplification by Southern blot, slot blot or FISH (Fluorescence *in situ* hybridization) as well as p185 expression by immunohistochemistry (IHC) [13, 14] in order to evaluate the prognostic value as well as the potential applications of *c-erbB-2* testing for predicting response of breast cancer to therapy. Despite the disparity of the results, it appeared that (i) the *c-erbB-2* alterations (mostly amplification) were correlated with a shortened time before disease relapse and a diminished overall survival and, (ii) that the most common association of p185 overexpression with therapeutic response was the apparent resistance of these tumors to hormone therapy alone. In summary, breast cancers overexpressing p185 correlate with poor prognostic factors and have a more aggressive behavior than those without overexpression.

Immunoenzymatic assay of p185 and evaluation of the degree of *c-erbB-2* amplification in breast cancers have rarely been performed simultaneously. When the quantitative method evaluating p185 was used to characterize *c-erbB-2* status of breast cancer samples, authors have realized a dichotomous distribution according to various cutoff values [16–19]. However, a dichotomous distribution leads to a great loss of information [20]. In this context, like Cuny et al. [21], we realized a subdivision into three subgroups (positive, intermediate and negative) according to two cutoff values (20 and 3 U/μg prot.). Such a categorization cannot be obtained when *c-erbB-2* amplification is used and therefore with p185 quantification it was possible to define a subset of breast cancers with moderate levels of p185. In our study, this subgroup of breast cancers (52.8% of cases) showed an ambiguous behavior as it was simultaneously associated with favourable prognostic factors (estrogen and progesterone receptors) and unfavourable ones (cathepsin D and axillary nodal involvement). Dittadi et al. [22] showed in 115 breast cancer specimens, that p185 evaluation by ELISA allowed definition of an intermediate concentration group with a significantly better disease-free survival than the low and high concentration groups. Therefore, it appears that a p185 intermediate expression in breast cancers is preferably associated with favourable outcome.

In the breast cancer population studied, we showed that the mean ER concentrations were different according to the p185 subgroup concerned. It has been observed that 17-β estradiol downregulates *c-erbB-2* oncogene expression [23, 24]. Thus, the high ER levels (249 ± 65 fmol/mg prot.) observed in the p185 intermediate subgroup might explain the moderate expression of p185. Conversely, the relatively low ER levels of p185 positive subgroup could be insufficient to repress *c-erbB-2* expression. On the other hand, the p185 negative subgroup presents relatively low ER levels and this fact is not in agreement with the former hypothesis. This observation suggests the intervention of other factors in the regulation of *c-erbB-2* expression. In addition, this subset of breast cancers might have negative prognostic significance as suggested by Koscielny et al. [25]. The antiestrogen tamoxifen upregulates *c-erbB-2* oncogene in estrogen-responsive breast cancers [26]. Thus, the absence of clinical response to antiestrogen therapy in 40% of patients with ER-positive disease might be due to the upregulation

c-erbB-2 oncogene in the primary tumor. Conversely, some ER-positive tumors are characterized by the absence of p185 expression and this suggests that tamoxifen and estradiol could control *c-erbB-2* expression only when this gene is switched on. The mechanism responsible for this switching on is not well known but it might be independent of the presence of estrogen receptors or dependent on the presence of other activated factors.

As it was observed by Cuny et al. [21], breast carcinomas of the lobular type rarely overexpress p185 and if so, only at intermediate levels, and this confirms the specific expression of p185 by the ductal type. Moreover, the absence of correlation of p185 expression with the histologic grade suggests an early implication of *c-erbB-2* alterations in the tumoral growth although some studies reported a correlation of p185 overexpression with high grade breast cancers [16, 27, 28]. The reasons for these discrepancies may be related to the mode of selection of the sample population, technical problems or the criteria used to define overexpression.

The paradoxical correlations of the p185 intermediate subgroup with the clinical and biological parameters suggests the existence of underlying factors which might discriminate the breast cancers having a truly pejorative prognosis from those having a more favourable issue. Christianson et al. [29] have identified, in human breast carcinomas, a NH2-terminal-truncated HER-2/neu protein of 95 kDa (namely p95HER-2/neu) which was expressed in 22.4% of breast cancers and was preferentially found in p185HER-2/neu-positive patients with lymph node involvement. A biological explanation of their findings was that loss of the extracellular domain of p185 from the p95 kinase combined with an amplified p185 signal in primary breast tumor cells could promote their metastasis to the lymph nodes.

There is at present no specific ligand known to bind p185 and p185 homodimerizes or heterodimerizes with the other members of the c-erbB family [30]. Siegel et al. [31] have shown, in transgenic mice, that coexpression of *c-erbB-2* and c-erbB-3 induced multiple mammary tumors that frequently metastasized to the lung. Moreover, they detected in human breast cancers an alternatively spliced form of *c-erbB-2* which encodes an ErbB-2 receptor harboring an in-frame deletion of 16 amino acids in the extracellular domain and can transform Rat-1 fibroblasts. Their observations infer that coexpression of ErbB-2 and ErbB-3 may play a critical role in the induction of human breast cancers, and evoke the possibility that activating mutations in the ErbB-2 receptor may also contribute to this process.

Here, we showed that quantification of p185HER-2/neu allows characterization of subsets of breast cancers which are, or are not, correlated with *c-erbB-2* amplification but which support divergent correlations with the classical prognostic factors. It appears that, considering the multiple alterations affecting the *c-erbB-2* oncogene (amplification, mutation, alternative splicing, overexpression) single determination of p185 expression by an ELISA method is well adapted to characterization of the *c-erbB-2* status of breast cancers and could improve clinical information. However, further investigations of the c-erbB family members could be useful for better characterization of breast cancers and particularly those harboring an overexpression of p185 without *c-erbB-2* gene amplification.

Acknowledgements

The excellent technical assistance of Ms. E. Chezy, C. Colombain, C. Vial and C. Ferniot is acknowledged. This work was supported by grants from the Institut National de la Santé et de la Recherche Médicale and the Ligue contre le cancer.

References

1. Bargmann CI, Hung MC, Weinberg RA: The *neu* oncogene encodes an epidermal growth factor receptor-related protein. Nature 319: 226–230, 1986a
2. Coussens L, Yang-Feng TL, Liao YC, Chen E, Gray A, McGrath J, Seeburg PH, Libermann TA, Schlessinger J, Francke U, Levinson A, Ullrich A: Tyrosine kinase receptor with extensive homology to EGF receptors shares chromosomal location with neu oncogene. Science (Washington DC) 230: 1132–1139, 1985
3. Rajkumar T, Gullick WJ: The type I growth factor receptors in human breast cancer. Breast Cancer Res Treat 29: 3–9, 1994
4. Shih C, Padhy LC, Murray M, Weinberg RA: Transforming genes of carcinomas and neuroblastomas induced into mouse fibroblasts. Nature 290: 261–264, 1981
5. Padhy LC, Shih C, Cowing D, Finkelstein R, Weinberg RA: Identification of a phosphoprotein specifically induced by the transforming DNA of rat neuroblastomas. Cell 28: 865–871, 1982
6. Bargmann CI, Hung MC, Weinberg RA: Multiple independent activations of the *neu* oncogene by a point mutation altering the transmembrane domain. Cell 45: 649–657, 1986
7. Bargmann CI, Weinberg RA: Increased tyrosine kinase activity associated with the protein encoded by the activated neu oncogene. Proc Natl Acad Sci USA 85: 5394–5398, 1988

8. Yarden Y: Agonistic antibodies stimulates the kinase encoded by the *neu* protoncogene in living cells but the oncogenic mutant is constitutively active. Proc Natl Acad Sci USA 87: 2569–2573, 1990
9. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, Levin WJ, Stuart SG, Udove J, Ullrich A: Studies of the *HER-2/neu* proto-oncogene in human breast and ovarian cancer. Science (Washington DC) 244: 707–712, 1989
10. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL: Human breast cancer: correlation of relapse and survival with the amplification of the *HER-2/neu* oncogene. Science (Washington DC) 235: 177–182, 1987
11. Descotes F, Pavy JJ, Adessi GL: Human breast cancer: correlation study between *HER-2/neu* amplification and prognostic factors in an unselected population. Anticancer Res 13: 119–124, 1993a
12. Descotes F, Pavy JJ, Adessi GL: *HER-2/neu* amplification in human breast cancer: Southern or Slot blotting amplification analysis? Eur J Cancer 29A: 650, 1993b
13. Révillion F, Bonneterre J, Peyrat JP: *ERBB2* oncogene in human breast cancer and its clinical significance. Eur J Cancer 34: 791–808, 1998
14. Ross JS, Fletcher JA: *HER-2/neu* (*c-erbB-2*) gene and protein in breast cancer. Am J Clin Pathol 112: S53–S67, 1999
15. Duffy MJ, Reilly D, Brouillet JP, McDermott EW, Faul C, O'Higgins N, Fennelly JJ, Maudelonde T, Rochefort H: Cathepsin D concentration in breast cancer cytosols: correlation with disease-free interval and overall survival. Clin Chem 38: 2114–2116, 1992
16. Dalifard I, Daver A, Goussard J, Lorimier G, Gosse-Brun S, Lortholary A, Larra F: p185 overexpression in 220 samples of breast cancer undergoing primary surgery: comparison with *c-erbB-2* gene amplification. Int J Molec Med 1: 855–861, 1998
17. Dittadi R, Donisi PM, Brazzale A, Marconato R, Spina M, Gion M: Immunoenzymatic assay of erbB-2 protein in cancer and non-malignant breast tissue. Relationships with clinical and biochemical parameters. Anticancer Res 12: 2005–2010, 1992
18. Marsigliante S, Muscella A, Ciardo V, Barker S, Leo G, Baker V, Mottaghi A, Vinson GP, Storelli C: Enzyme-linked immunosorbent assay of *HER-2/neu* gene product (p185) in breast cancer: its correlation with sex steroid receptors, cathepsin D and histologic grades. Cancer Lett 75: 195–206, 1993
19. Valeron PF, Chirino R, Vega V, Falcon O, Rivero JF, Torres S, Leon L, Fernandez L, Pestano J, Diaz-Chico B, Diaz-Chico JC: Quantitative analysis of p185(HER-2/neu) protein in breast cancer and its association with other prognostic factors. Int J Cancer 74: 175–179, 1997
20. Altman DG, Lausen B, Sauerbrei W, Schumacher M: Danger of using 'optimal' cutpoints in the evaluation of prognostic factors. J Natl Cancer Inst 86: 829–835, 1994
21. Cuny M, Simony-Lafontaine J, Rouanet P, Grenier J, Valles H, Lavaill R, Louason G, Causse A, Lequeux N, Thierry C et al.: Quantification of ERBB2 protein expression in breast cancer: three levels of expression defined by their clinico-pathological correlations. Oncol Res 6: 169–176, 1994
22. Dittadi R, Brazzale A, Pappagallo G, Salbe C, Nascimben O, Rosabian A, Gion M: ErbB2 assay in breast cancer: possibly improved clinical information using a quantitative method. Anticancer Res 17: 1245–1247, 1997
23. Read LD, Keith D Jr, Slamon DJ, Katzenellenbogen BS: Hormonal modulation of *HER-2/neu* protooncogene messenger ribonucleic acid and p185 protein expression in human breast cancer cell lines. Cancer Res 50: 3947–3951, 1990
24. Russel KS, Hung MC: Transcriptional repression of the *neu* protooncogene by estrogen stimulated estrogen receptor. Cancer Res 52: 6624–6629, 1992
25. Koscielny S, Terrier P, Daver A, Wafflart J, Goussard J, Ricolleau G, Delvincourt C, Delarue JC: Quantitative determination of *c-erbB-2* in human breast tumours: potential prognostic significance of low values. Eur J Cancer 34: 476–481, 1998
26. Antoniotti S, Maggiora P, Dati C, De Bortoli M: Tamoxifen up-regulates *c-erbB-2* expression in oestrogen-responsive breast cancer cells *in vitro*. Eur J Cancer 28: 318–321, 1992
27. Barnes DM, Lammie GA, Millis RR, Gullick WL, Allend S, Altman DG: An immunohistochemical evaluation of *c-erbB-2* expression in human breast carcinoma. Br J Cancer 58: 448–452, 1988
28. Schroeter CA, De Potter CR, Rasthsmann K, Willighagen RGJ, Greep JC: *c-erbB-2* positive tumours behave more aggressively in the first years after diagnosis. Br J Cancer 60: 426–429, 1992
29. Christianson TA, Doherty JK, Lin YJ, Ramsey EE, Holmes R, Keenan EJ, Clinton GM: NH2-terminally truncated *HER-2/neu* protein: relationship with shedding of the extracellular domain and with prognostic factors in breast cancer. Cancer Res 58: 5123–5129, 1998
30. Klapper LN, Glathe S, Vaisman N, Hynes NE, Andrews GC, Sela M, Yarden Y: The ErbB-2/HER2 oncoprotein of human carcinomas may function solely as a shared coreceptor for multiple stroma-derived growth factors. Proc Natl Acad Sci USA 96: 4995–5000, 1999
31. Siegel PM, Ryan ED, Cardiff RD, Muller WJ: Elevated expression of activated forms of Neu/ErbB-2 and ErbB-3 are involved in the induction of mammary tumors in transgenic mice: implications for human breast cancer. EMBO J 18: 2149–2164, 1999

Address for offprints and correspondence: Gérard L. Adessi, Institut d'Etude et de Transfert de Génes, 240 route de Dole, 25000 Besançon, France; *Tel.:* +33.3.81.52.33.00; *Fax:* +33.3.81.52.64.80; *E-mail:* gerard.adessi@ufc-chu.univ-fcomte.fr

[CANCER RESEARCH 62, 2625–2629, May 1, 2002]

Targets of Gene Amplification and Overexpression at 17q in Gastric Cancer[1]

Asta Varis,[2] Maija Wolf,[2] Outi Monni,[2] Marja-Leena Vakkari, Arto Kokkola, Chris Moskaluk, Henry Frierson, Jr., Steven M. Powell, Sakari Knuutila, Anne Kallioniemi, and Wa'el El-Rifai[3]

Department of Medical Genetics [A. V., M-L. V., S. K.], Biomedicum Biochip Center [O. M.], and Department of Surgery [A. K.], University of Helsinki and Helsinki University Central Hospital, Helsinki 00290, Finland; Cancer Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, Maryland 20892 [M. W., O. M., A. K.]; Laboratory of Cancer Genetics, University of Tampere, Tampere 33014, Finland [A. K.]; and Department of Pathology [C. M., H. F.] and Digestive Health Center of Excellence [S. M. P., W. E-R.], University of Virginia Health System, Charlottesville, Virginia 22908-0708

ABSTRACT

DNA copy number gains and amplifications at 17q are frequent in gastric cancer, yet systematic analyses of the 17q amplicon have not been performed. In this study, we carried out a comprehensive analysis of copy number and expression levels of 636 chromosome 17-specific genes in gastric cancer by using a custom-made chromosome 17-specific cDNA microarray. Analysis of DNA copy number changes by comparative genomic hybridization on cDNA microarray revealed increased copy numbers of 11 known genes (*ERBB2, TOP2A, GRB7, ACLY, PIP5K2B, MPRL45, MKP-L, LHX1, MLN51, MLN64,* and *RPL27*) and seven expressed sequence tags (ESTs) that mapped to 17q12-q21 region. To investigate the genes transcribed at the 17q, we performed gene expression analyses on an identical cDNA microarray. Our expression analysis showed overexpression of 8 genes (*ERBB2, TOP2A, GRB2, AOC3, AP2B1, KRT14, JUP,* and *ITGA3*) and two ESTs. Of the commonly amplified transcripts, an uncharacterized EST AA552509 and the *TOP2A* gene were most frequently overexpressed in 82% of the samples. Additional studies will be initiated to understand the possible biological and clinical significance of these genes in gastric cancer development and progression.

INTRODUCTION

Gastric carcinoma is one of the most common malignancies worldwide and is the second most frequent cause of cancer-related death (1). Moreover, cardia, gastroesophageal junction, and esophageal adenocarcinomas have the most rapidly rising incidence of all visceral malignancies in the United States and Western world for reasons that are unclear (2). Previous studies have documented the importance of genetic alterations affecting known oncogenes, tumor suppressor genes, and mismatch repair genes in the development of gastric cancer (3, 4). Several genes, such as *cMET, ERBB2, MYC,* and *MDM2,* are amplified in 10–25% of tumors, and their amplification is associated with advanced disease (3, 5). Comprehensive DNA copy number analyses of gastric cancers using CGH[4] have demonstrated recurrent DNA copy number changes on several chromosomal regions. Gains at 17q have been shown to be frequent in gastric cancers (6). However, the critical regions of genetic alterations are large, and the target genes for amplification at 17q remain unknown.

Characterization of the chromosomal regions involved in DNA copy number changes is likely to reveal genes important for the development of gastric cancer. In the present study, we used a custom-made chromosome 17-specific cDNA microarray to systematically evaluate copy numbers and expression levels of genes at 17q in gastric carcinomas.

MATERIALS AND METHODS

Samples. Sixteen gastric cancer xenografts, four gastric cancer cell lines (CRL-5822, CRL-5974, CRL-5973, and CRL-1739) from the American Type Culture Collection (Manassas, VA), and five primary gastric cancers were used in this study. The cell line (CRL-1739) with normal DNA copy number of chromosome 17 was included as a control in Northern blot hybridizations. The cell lines were cultured under recommended conditions. Xenografting of gastric cancers was performed as described earlier (7). All tumors included in this study were dissected and verified histologically to be composed predominantly of neoplastic tissues. We have earlier characterized the DNA copy numbers of the cell lines and xenografts using "chromosomal" CGH. The details of the DNA copy numbers of the xenografts have been reported elsewhere (7). Fig. 1*A* summarizes the chromosomal CGH results for chromosome 17.

Chromosome 17-specific cDNA Microarray. The construction of the chromosome 17-specific cDNA microarray has been described previously (8). Briefly, the cDNA microarray contained a total of 636 clones, including 88 house keeping genes, 201 known genes from chromosome 17, and 435 EST clones from radiation hybrid map intervals D17S933-D17S930 (293–325 cR, the 17q12-q21 region) and D17S791-D17S795 (333–435 cR, the 17q23-q24 region). The preparation and printing of the cDNA clones on glass slides were performed as described elsewhere (9).

Copy Number and Expression Analyses by cDNA Microarrays. Genomic DNA was extracted from eight xenografts (X11, X27, X57, X71, X75, X79, X83, and X95) and three cell lines (CRL-5822, CRL-5973, and CRL-5974). All cases had gains or high-level amplification at 17q by chromosomal CGH (Fig. 1). Normal genomic DNA was used as a reference in all experiments. Copy number analysis using CGH microarray was performed as described previously (8, 10). Briefly, 20 μg of genomic DNA were digested for 14–18 h with *Alu*I and *Rsa*I restriction enzymes (Life Technologies, Inc., Rockville, MD) and purified by phenol/chloroform extraction. Digested gastric cancer test DNA (6 μg) was labeled with Cy3-dUTP (Amersham Pharmacia Biotech, Piscataway, NJ) and 6 μg of reference DNA with Cy5-dUTP using Bioprime Labeling kit (Life Technologies, Inc.). Hybridization was done according to the protocol by Pollack *et al.* (10) and posthybridization washes as described previously (11).

Total RNA was extracted from eight xenografts (X43, X49, X57, X68, X75, X76, X80, and X95) and three gastric cancer cell lines (CRL-5822, CRL-5973, and CRL-5974) by using RNeasy kit (Qiagen, GmbH, Hilden, Germany). A pool of four normal gastric epithelial tissue samples, enriched for the epithelial layer of the stomach through dissection and mucosal scrapping, was used as a standard reference in all experiments. Reference RNA (100 μg) was labeled with Cy5-dUTP and 80 μg of test RNA with Cy3-dUTP by use of oligodeoxythymidylate-primed polymerization by SuperScript II reverse transcriptase (Life Technologies, Inc.). The labeled cDNAs were hybridized on microarrays as described previously (11, 12).

For both the copy number and expression analyses, the fluorescence intensities at the cDNA targets were measured by using a laser confocal scanner (Agilent Technologies, Palo Alto, CA). The fluorescent images from the test and control hybridizations were scanned separately, and the data were analyzed using the DEARRAY software (13). After the subtraction of background intensities, the average intensities of each spot in the test hybridization were divided by the average intensity of the corresponding spot in the control hybridization. On the basis of our earlier reports (8, 14), clones that showed a

Received 12/15/01; accepted 2/27/02.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported in part by scholarships from the Biomedicum Foundation and Paulo Research Foundation in Finland and by grants from the University of Virginia Cancer Center and the American Cancer Society Grant 5-37408 sub.

[2] A. V., M. W., and O. M. contributed equally to this manuscript.

[3] To whom requests for reprints should be addressed, at the Digestive Health Center of Excellence, University of Virginia Health System, P.O. Box 800708, Charlottesville, VA 22908-0708. Phone: (434) 243-6158; Fax: (434) 243-6169; E-mail: wme8n@virginia.edu.

[4] The abbreviations used are: CGH, comparative genomic hybridization; EST, expressed sequence tag; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; RT-PCR, reverse transcription-PCR; AP, adapter protein.

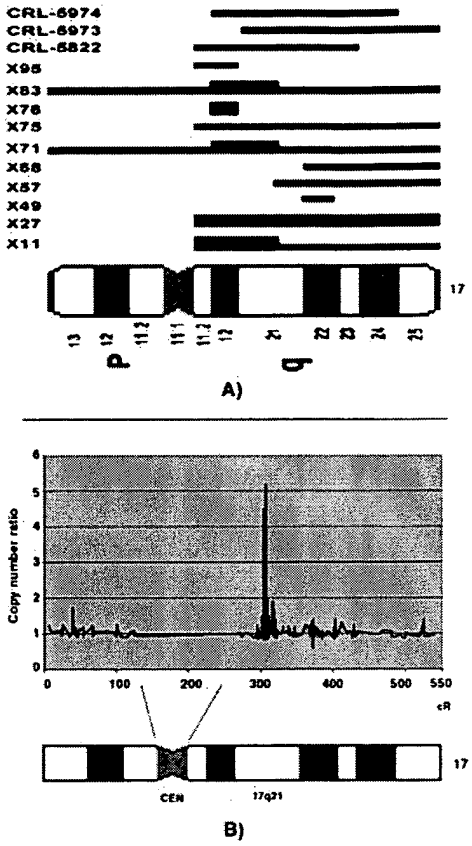


Fig. 1. DNA copy numbers in gastric cancer. *A*, summary of gains and high-level amplifications affecting chromosome 17 in gastric cancer xenografts and cell lines by chromosomal CGH. *Horizontal bars*, the extent of the copy number aberration in each sample. High-level amplifications are presented as *wide bars*. xenograft samples X43, X80, and cell line CRL-1739 had no detectable gains on chromosome 17. *B*, copy numbers survey of chromosome 17-specific genes in X83 gastric cancer xenograft by CGH microarray. The copy number ratios were plotted as a function of the position of the clones in the radiation hybrid map in cR scale. Individual data points were connected with a *line*. The chromosome 17 ideogram is shown *below* for visual comparison only.

copy number ratio ≥ 1.5 were considered as amplified, and clones that showed an expression ratio ≥ 3 were considered as overexpressed. Clones that showed such increased ratios in the self *versus* self control experiment were excluded from the analysis.

Northern Hybridization. Total RNA was extracted from four gastric cancer cell lines and two normal stomach specimens using the RNeasy kit (Qiagen, GmbH). The Northern hybridization was performed using standard methods. Briefly, 10 μg of total RNA were size-fractionated on a 1% agarose gel containing formaldehyde and transferred on a Nytran membrane (Schleicher & Schuel, Keene, NH). The membrane was prehybridized for 1 h at 65°C in Express hybridization solution (Clontech, Palo Alto, CA) together with sheared Herring sperm DNA (10 μg/ml; Research Genetics, Huntsville, AL). Sequence-verified cDNA inserts were labeled with P^{32} by random priming (Prime-It; Stratagene, La Jolla, Ca). Hybridization was performed in the Express hybridization solution (Clontech) at 65°C overnight followed by washes in 2 × SCC/SDS solutions. Signals were detected by autoradiography. The normal gastric tissues and CRL-1739 cell line (normal chromosome 17 on CGH) were used as control samples. A GAPDH cDNA was used as a control probe.

Multiplex RT-PCR. Multiplex RT-PCR was used to validate the cDNA array results for the two most overexpressed genes (*ESTAA552509* and *TOP2A*) using seven xenografted and six primary gastric cancer samples. For reference expression, a pool of normal gastric epithelial tissues obtained from different individuals was used. Primary tumors of four xenografts were included in the analyses. mRNA was purified from the tissues using mRNeasy (Qiagen), and cDNA synthesis was performed using Advantage RT-for-PCR Kit (Clontech). In each PCR reaction, primers for the human *GAPDH* gene were used as an internal reference. The PCR reactions were done using standard protocol for 28 cycles. We confirmed the reproducibility of the method by repeating the RT-PCR twice, and the results were consistent. The primers used for the RT-PCR were obtained from GeneLink (Hawthorne, NY), and their sequences are available on request.

RESULTS

Detailed Characterization of the 17q Amplification Using Chromosome-specific Microarray. Copy number levels of 636 chromosome 17-specific genes were evaluated by CGH microarray in eight xenografts (X11, X27, X57, X71, X75, X79, X83, and X95) and three gastric cancer cell lines (CRL-5822, CRL-5973, and CRL-5974) that

Table 1 *Summary of copy number ratios of 18 chromosome 17q12-q21 transcripts in eight xenografts and three cell lines of gastric cancer by CGH microarray*[a]

Gene	Unigene Id	Accession	Alignment[b]	Locus[b]	Samples										
					X11	X27	X57	X71	X75	X79	X83	X95	CRL-5822	CRL-5973	CRL-5974
MRPL45 Mitochondrial ribosomal protein L45	Hs. 19347	AI277785	38274220/51922787	17q12/17q21.3	1.4	**1.5**	**1.1**	1	**2.4**	1.3	**1.6**	**1.6**	**2**	1.4	1.2
MKP-1 like protein tyrosine phosphatase (MKP-L)	Hs. 91448	AA129677	38747279	17q12	**1.9**	**1.6**	0.8	1.4	1.1	1.1	1.2	**2.8**	1.4	1.4	**1.7**
LIM homeobox protein 1 (LHX1)	Hs. 157449	AI375565	39307916	17q12	1	1.3	**2.1**	1.1	**2.8**	1.2	1	**2.8**	1.4	**2.5**	**2.1**
Phosphatidylinositol-4-phosphate 5-kinase, type II, β (PIP5K2B)	Hs. 6335	H80263	40617731	17q12	1	0.9	**1.6**	**1.7**	1.3	1.3	0.9	**1.7**	1.4	**2.8**	0.8
EST	Hs. 91668	H16094	41911205	17q21.1	1.2	0.8	1.2	1	1.4	**3.8**	**4.7**	0.9	**3.1**	1.2	0.8
EST (FLJ20940 hypothetical protein)	Hs. 286192	AA552509	41868584	17q21.1	1.4	1.1	**1.7**	1.2	1.4	**5.8**	**7.3**	**1.6**	**1.5**	1.4	1.1
H. sapiens MLN64 mRNA	Hs. 77628	AA504615	41877246	17q21.1	1.1	1.1	1.2	1.2	1.2	**5.5**	**4.2**	1.1	**3.1**	1.4	1.1
V-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2 (ERBB2)	Hs. 323910	AA446928	41940229	17q21.1	1	1	1	1	1.1	**2.6**	**2.1**	1	**1.7**	1.1	1
EST	Hs. 46645	AA283905	41972680	17q21.1	1.3	1.1	1	1.3	0.9	**5.1**	**10.4**	1.3	**2**	0.8	1.3
EST	Hs. 318893	AA455291	41978415	17q21.1	1.1	0.9	1.1	1	0.9	**7.8**	**2.5**	0.8	**3.2**	1.4	0.9
Growth factor receptor-bound protein 7 (GRB7)	Hs. 86859	H53703	41989210	17q21.1	1.4	1	1	1.1	1	**5.7**	**8.9**	0.9	**2.7**	1.4	1
H. sapiens MLN51 mRNA	Hs. 83422	R52974	42331857	17q21.1	**1.6**	1.2	1.1	**1.6**	0.9	**1.9**	**1.9**	**1.5**	1.4	**2.1**	1.1
Topoisomerase (DNA) II α (170kD) (TOP2A)	Hs. 156346	AA026682	42521254	17q21.2	1.3	1.1	1.2	**1.6**	1	**1.7**	**1.6**	1.4	**1.6**	1.4	1.1
EST	Hs. 13268	AA514361	44056922	17q21.2	1.2	1.1	1	**1.9**	1.3	**2.3**	1.2	**1.8**	**1.5**	**1.8**	**1.5**
ATP citrate lyase (ACLY)	Hs. 174140	R55974	44075311	17q21.2	1.2	1	1	**1.6**	1	**1.6**	1.1	**1.6**	1.4	**1.6**	**2.3**
EST	Hs. 38039	H62271	44574937	17q21.2	**5.7**	**2.9**	**2.3**	**6.7**	**3.1**	1	**1.7**	**2.5**	**3**	**4.3**	**2.7**
Ribosomal protein L27 (RPL27)	Hs. 111611	AA190881	45301136	17q21.2	**4.6**	**2.5**	**2.2**	**6.1**	**1.7**	0.9	**2**	2	**2.1**	1.4	**2.9**
EST (DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 8, DDX8)	Hs. 171872	AI540663	46054957	17q21.3	1.4	**1.7**	0.5	1.4	0.9	**2.9**	**1.6**	1.2	**1.9**	**3.7**	1.1

[a] Copy number ratios above the 1.5 threshold are shown in bold.
[b] Alignment (bp position) and locus are shown according to Santa Cruz August freeze 2001 assembly.

Table 2 *Summary of expression levels of 10 chromosome 17q12-q21 transcripts in eight xenografts and three cell lines of gastric cancer by cDNA microarray*[a]

Gene	Unigene Id	Accession	Alignment[b]	Locus[b]	Samples										
					X43	X49	X57	X68	X75	X76	X80	X95	CRL-5822	CRL-5973	CRL-5974
Adaptor-related protein complex 2, β1 subunit (AP2B1)	Hs. 74626	H29927	37327700	17q12	1.1	1	2.9	**3.1**	0.9	1.1	0.9	0.8	**4.8**	**7.4**	**4.6**
EST (Hypothetical protein FLJ20940)	Hs. 286192	AA552509	41868584	17q21.1	**21.9**	**4.5**	**6.4**	**7.6**	**10.2**	**10**	**17.3**	0.6	**12.1**	**11.2**	0.6
V-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2 (ERBB2)	Hs. 323910	AA446928	41940229	17q21.1	1	1	**3**	**3.7**	1.4	0.9	1.3	0.7	**24.6**	0.7	1
Topoisomerase (DNA) II α (170kD) (TOP2A)	Hs. 156346	AA026682	42521254	17q21.2	**4.1**	**6.1**	**16**	**4.5**	2.8	**6.6**	**3**	1.4	**5.6**	**7.6**	**6.8**
Keratin 14 (KRT14)	Hs. 117729	H44127	43757143	17q21.2	**3.9**	1.4	1.1	1.6	**3.5**	1	1.2	0.6	**3.8**	1.3	0.8
Junction plakoglobin (JUP)	Hs. 2340	R06417	43994962	17q21.2	**3.1**	2.8	0.9	1.2	**4.3**	0.9	2.6	**3.4**	**5**	1.9	2.5
Amine oxidase, copper containing 3 (AOC3)	Hs. 198241	T77398	45078066	17q21.2	**4.6**	1.9	**4.5**	2.6	2.1	2.6	**4.2**	1.6	**3.1**	1.8	**5**
Integrin, α-3 (ITGA3)	Hs. 265829	AA424695	54688140	17q21.3	**4.8**	1.5	0.9	**4.5**	**3.4**	1.3	1.1	1.2	2.7	2.2	0.5
EST	Hs. 56105	AA284262	65817334	17q23.2	1	1.7	**5.2**	**3**	**15.6**	2.8	0.6	1.9	1.3	0.6	0.4
Growth factor receptor-bound protein 2 (GRB2)	Hs. 296381	AA449831	81840742	17q25.1	0.8	0.8	2.2	1.4	1.3	1.8	1.1	1.3	**3.1**	**5.6**	**3.5**

[a] Expression ratios above the 3 threshold are shown in bold.
[b] Alignment (bp position) and locus are shown according to Santa Cruz August freeze 2001 assembly.

showed gain or high-level amplification affecting chromosome 17 by chromosomal CGH (Fig. 1). CGH microarray analysis revealed increased DNA copy numbers (ratio ≥ 1.5) in three or more cases for 11 genes and seven ESTs that map to 17q12 (4 clones) and 17q21 (14 clones; Table 1). The amplified genes/ESTs were localized at 302–321 cR in the radiation hybrid map[5] (Fig. 1*B*) and between 38274220–46054957 bp at 17q, according to the University of California Santa Cruz's August freeze 2001 assembly of the human genome sequence.[6] The two most consistently amplified clones were EST (H62271) and ribosomal protein L27 (82%). Other frequently amplified genes included *TOP2A, EST AA552509,* and *ERBB2*. The details of the copy numbers and location of these genes/ESTs are listed in Table 1.

Gene Expression Profiling of 17q Using cDNA Microarrays. Parallel expression survey in eight xenografts (X43, X49, X57, X68, X75, X76, X80, and X95) and the three cell lines identified 10 transcripts at 17q whose expression was elevated (ratio ≥ 3) in at least three specimens, as compared with the normal gastric epithelial cells (Table 2; Fig. 2). Three of the commonly amplified sequences (*TOP2A, ERBB2,* and EST AA552509) that map to 17q21 were also overexpressed frequently in our cDNA expression analyses. The two most consistently affected transcripts were EST AA552509 (82%) and the *TOP2A* (82%).

Other frequently overexpressed genes included *AOC3* (45%), *JUP* (36%), *ERBB2* (27%), *ITAG3* (27%), and *KRT14* (27%) at 17q21 region, as well as *AP2B1* at 17q12, EST AA284262 at 17q23, and *GRB2* at 17q25 (Table 2; Fig. 2).

Northern Blotting. Northern analysis was used as an independent expression assay to validate the cDNA microarray results. Because of the limited availability of RNA from the xenografted tumors, only cell lines were analyzed. Three genes, *EST AA552509, TOP2A,* and *ERBB2,* that showed overexpression in one or more cell lines by cDNA microarray were selected for analysis. Results from the Northern analysis confirmed the cDNA microarray data. *ERBB2* was highly overexpressed in CRL-5822 cell line, *TOP2A* in all three cell lines, and EST AA552509 in CRL-5822 and CRL-5973 (Fig. 3). These genes were not expressed in the normal gastric epithelial sample or the gastric cell line (CRL-1739) that had normal chromosome 17 DNA copy numbers by chromosomal CGH (Fig. 3).

Multiplex RT-PCR. Expression analyses with RT-PCR showed elevated expression of *TOP2A* and EST AA552509 in all tested tumor samples, whereas no expression was seen in the pool of normal gastric epithelial tissues (Fig. 3). The xenografts and their corresponding primaries showed similar levels of expression.

DISCUSSION

Studies by chromosomal CGH have indicated that 17q is amplified frequently in gastric cancer. Here we used a custom-made cDNA

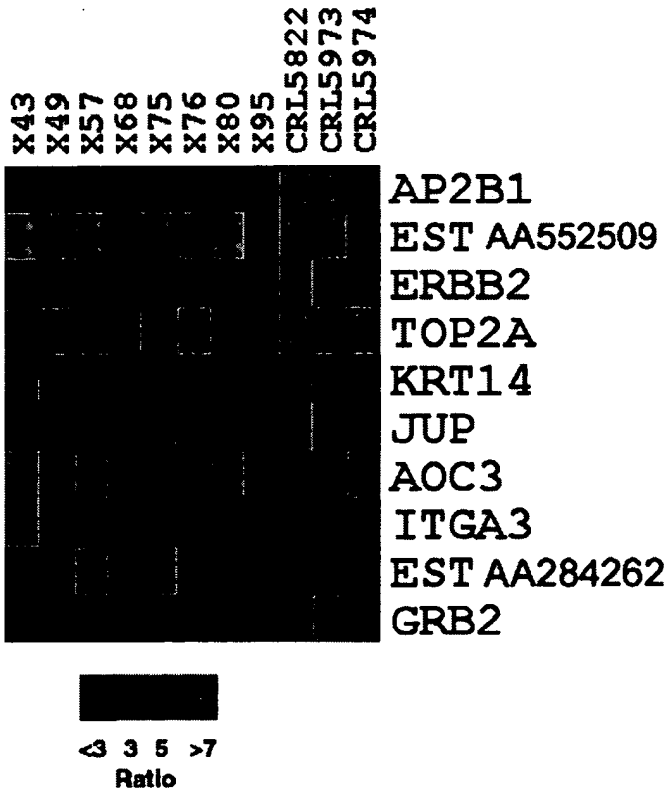


Fig. 2. Expression patterns of the most commonly overexpressed genes in gastric cancer xenografts and cell lines. Names of the genes are indicated on the *right*. Color coding for the expression ratios is shown *below* the graph. This image was created using Tree view software written by Michael Eisen, copyright 1998–1999, Stanford University.

[5] Internet address: http://www.ncbi.nlm.nih.gov/genemap.
[6] Internet address: http://genome.ucsc.edu.

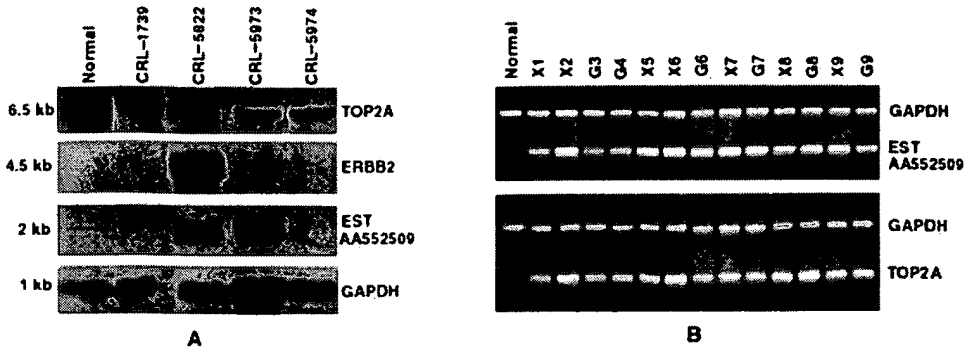


Fig. 3. Validation of overexpressed genes in gastric cancer. *A*, Northern analysis of *TOP2A, ERBB2,* and EST AA552509 expression in normal gastric tissue and four gastric cancer cell lines. CRL-1739 had normal copy numbers by CGH. The size of each transcript is indicated on the *right* side of the corresponding picture. *GAPDH* was used as a loading control. *B*, expression analysis by multiplex RT-PCR in normal gastric tissue, seven xenografts (indicated by X-number), and six primary gastric cancers (indicated by G-number). Xenografts and their corresponding primary cancers have the same number. The names of the gene are shown on the *right*.

microarray that contained 636 cDNA clones from chromosome 17 to systematically analyze the copy number changes at 17q in eight gastric cancer xenografts and three cell lines. The CGH microarray analyses showed increased copy number ratios for 18 clones that were localized to the 17q12-q21 region. To identify those genes that are activated through increased copy number, we performed a comprehensive gene expression profiling using the same chromosome 17-specific cDNA microarray. Three of the commonly amplified transcripts (*TOP2A, ERBB2,* and EST AA552509) that map to 17q21 were overexpressed frequently in our analyses and might, therefore, represent putative amplification target genes in gastric cancer. The cDNA microarray results were validated using Northern and RT-PCR analyses.

The two most frequently overexpressed genes in our samples were the EST AA552509 and *TOP2A*. In addition, *ERBB2* was also amplified and overexpressed in >30% of tumors. Our data show that these genes are overexpressed in gastric cancers with no indication of their expression in normal gastric epithelial tissues. The overexpression of EST AA552509 has not been reported before and might be important for gastric carcinogenesis or have a possible value as a tumor marker or therapeutic target. On the other hand, the importance of *TOP2A,* and *ERBB2* in cancer, especially breast cancer, is well known (15, 16). *TOP2A* is an enzyme that catalyzes ATP-dependent strand-passing reactions and functions in DNA replication and chromosome condensation and segregation (17). *TOP2A* is a molecular target for many anticancer drugs (topo2 inhibitors). *ERBB2* is amplified frequently in breast cancer and has been shown to be an independent prognostic factor (18, 19). In breast cancer, *TOP2A* is often coamplified with *ERBB2* (20, 21). In our gastric adenocarcinomas, amplification and overexpression of *TOP2A* were independent of and also more frequent than *ERBB2*. Previous studies of *ERBB2* in gastric cancer have shown that the frequency of its overexpression varies from 9 to 38% (22, 23), which is in agreement with our findings. Our results provide additional evidence that clinical studies are required to determine the possibility that *TOP2A* and *ERBB2* are useful targets for cancer therapy in gastric cancer patients with these molecular alterations.

The up-regulation of *GRB2, JUP,* and *ITAG3* genes in the present study supports our earlier results that show these genes to be overexpressed in gastric cancer (7). Interestingly, studies in breast cancer suggest that *GRB2* may mediate transmission of *ERBB2* oncogenic signals, which in turn activate mitogen-activated protein kinase pathway (24, 25). *GRB2* is a widely expressed protein, which plays a crucial role in activation of several other growth factors (26).

KRT14, AOC3, and *AP2B1* were overexpressed in ≥3 of 11 of our gastric cancers. Copper-containing amino oxidases, such as *AOC3,* are involved in the catabolism of putrescine and histamine and are also involved in the regulation of growth and apoptosis (27). The AP2B1 is a member of AP complexes that function as vesicle coat components in different membrane traffic pathways. AP-2 complex associates with the plasma membrane and directs the internalization of trafficking cell surface protein (28). However, there is no information about the possible role of these genes in cancer.

Our study has identified genes that are coamplified at 17q12 and 17q21 amplicons that are not altered transcriptionally in comparison of tumors to normal reference samples. The lack of correlation between some amplified genes and their expression profile suggests that these genes are not critical targets at the 17q amplicon but might be coamplified together with critical genes within the amplicon structure. We also found genes that were overexpressed but not amplified by CGH microarrays. These results in CGH microarray may be attributed to the resolution of CGH-based technologies. On the other hand, upstream gene regulation and/or mutations are known as important biological mechanisms in transcriptional regulation irrespective of gene copy number.

Comparison of this gastric cancer study with our earlier data from breast cancer using the same cDNA microarray revealed a different pattern of alterations affecting chromosome 17 (8, 14). In breast cancer, two common regions of increased copy number and expression, 17q12-q21 and 17q23, were observed. In addition, the genes influenced by the 17q12-q21 amplification in gastric cancer differed from those in breast cancer where *ERBB2* was among the most strongly affected (8, 14). These results indicate that although 17q is involved frequently in copy number alterations in several cancers, the target loci and genes might be different from one tumor type to another.

In summary, the present study demonstrates that although the 17q region contains hundreds of genes, only three genes were frequently amplified and overexpressed in gastric cancers, as compared with normal gastric epithelial tissues. The consistent overexpression of *TOP2A* in our gastric cancers suggests that this gene may be a potential target for topo2 inhibitors in gastric cancer patients. The overexpression of EST AA552509, in the majority of our samples, suggests that this novel gene may play a critical role in gastric tumorigenesis. We have initiated additional studies to explore the possible biological and clinical significance of these genes in gastric cancer development and progression.

ACKNOWLEDGMENTS

We thank Jeffrey C. Harper for his technical assistance.

REFERENCES

1. Whelan, S., Parkin, D., and Masuyer, E. Trends in Cancer Incidence and Mortality. IARC Scientific Publ. No. 102. Lyon, France: IARC, 1993.
2. Locke, G. R. R., Talley, N. J., Carpenter, H. A., Harmsen, W. S., Zinsmeister, A. R., and Melton, L. J. R. Changes in the site- and histology-specific incidence of gastric cancer during a 50-year period. Gastroenterology, *109:* 1750–1756, 1995.
3. Werner, M., Becker, K. F., Keller, G., and Hofler, H. Gastric adenocarcinoma: pathomorphology and molecular pathology. J. Cancer Res. Clin Oncol., *127:* 207–216, 2001.

4. Lin, W., Kao, H. W., Robinson, D., Kung, H. J., Wu, C. W., and Chen, H. C. Tyrosine kinases and gastric cancer. Oncogene, *19:* 5680–5689, 2000.
5. Allgayer, H., Babic, R., Gruetzner, K. U., Tarabichi, A., Schildberg, F. W., and Heiss, M. M. c-erbB-2 is of independent prognostic relevance in gastric cancer and is associated with the expression of tumor-associated protease systems. J. Clin. Oncol., *18:* 2201–2209, 2000.
6. El-Rifai, W., Frierson, H. J., Moskaluk, C., Harper, J., Petroni, G., Bissonette, E., Knuutila, S., and Powell, S. Genetic differences between adenocarcinomas arising in Barrett's esophagus and gastric mucosa. Gastroenterology, *121:* 592–598, 2001.
7. El-Rifai, W., Frierson, H. J., Harper, J., Powell, S., and Knuutila, S. Expression profiling of gastric adenocarcinoma using cDNA array. Int. J. Cancer, *92:* 828–832, 2001.
8. Monni, O., Barlund, M., Mousses, S., Kononen, J., Sauter, G., Heiskanen, M., Paavola, P., Avela, K., Chen, Y., Bittner, M. L., and Kallioniemi, A. Comprehensive copy number and gene expression profiling of the 17q23 amplicon in human breast cancer. Proc. Natl. Acad. Sci. USA, *98:* 5711–5716, 2001.
9. DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A., and Trent, J. M. Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat. Genet., *14:* 457–460, 1996.
10. Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D., and Brown, P. O. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat. Genet., *23:* 41–46, 1999.
11. Mousses, S., Bittner, M., Chen, Y., Dougherty, E., Baxevanis, A., Meltzer, P., and Trent, J. (eds.). Functional Genomics, pp. 113–137. Oxford: Oxford University Press, 2000.
12. Khan, J., Simon, R., Bittner, M., Chen, Y., Leighton, S. B., Pohida, T., Smith, P. D., Jiang, Y., Gooden, G. C., Trent, J. M., and Meltzer, P. S. Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. Cancer Res., *58:* 5009–5013, 1998.
13. Chen, Y., Dougherty, E., and Bittner, M. Ratio-based decision and the quantitative analyses of cDNA microarray images. J. Biomed. Optics, *2:* 364–374, 1997.
14. Kauraniemi, P., Bärlund, M., Monni, O., and Kallioniemi, A. New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNA microarrays. Cancer Res., *61:* 8235–8240, 2001.
15. Harris, L. N., Yang, L., Liotcheva, V., Pauli, S., Iglehart, J. D., Colvin, O. M., and Hsieh, T. S. Induction of topoisomerase II activity after ErbB2 activation is associated with a differential response to breast cancer chemotherapy. Clin. Cancer Res., *7:* 1497–1504, 2001.
16. Cuello, M., Ettenberg, S. A., Clark, A. S., Keane, M. M., Posner, R. H., Nau, M. M., Dennis, P. A., and Lipkowitz, S. Down-regulation of the erbB-2 receptor by trastuzumab (herceptin) enhances tumor necrosis factor-related apoptosis-inducing ligand-mediated apoptosis in breast and ovarian cancer cell lines that overexpress erbB-2. Cancer Res., *61:* 4892–4900, 2001.
17. Osheroff, N. DNA topoisomerases. Biochim. Biophys. Acta, *1400:* 1–2, 1998.
18. Piccart, M., Lohrisch, C., Di Leo, A., and Larsimont, D. The predictive value of HER2 in breast cancer. Oncology, *61* (Suppl. S2)*:* 73–82, 2001.
19. Menard, S., Fortis, S., Castiglioni, F., Agresti, R., and Balsari, A. HER2 as a prognostic factor in breast cancer. Oncology, *61* (Suppl. S2)*:* 67–72, 2001.
20. Jarvinen, T. A., Kononen, J., Pelto-Huikko, M., and Isola, J. Expression of topoisomerase IIα is associated with rapid cell proliferation, aneuploidy, and c-erbB2 overexpression in breast cancer. Am. J. Pathol., *148:* 2073–2082, 1996.
21. Smith, K., Houlbrook, S., Greenall, M., Carmichael, J., and Harris, A. L. Topoisomerase II α co-amplification with erbB2 in human primary breast cancer and breast cancer cell lines: relationship to m-AMSA and mitoxantrone sensitivity. Oncogene, *8:* 933–938, 1993.
22. Ross, J. S., and McKenna, B. J. The HER-2/*neu* oncogene in tumors of the gastrointestinal tract. Cancer Investig., *19:* 554–568, 2001.
23. Gurel, S., Dolar, E., Yerci, O., Samli, B., Ozturk, H., Nak, S. G., Gulten, M., and Memik, F. The relationship between c-erbB-2 oncogene expression and clinicopathological factors in gastric cancer. J. Int. Med. Res., *27:* 74–78, 1999.
24. Dankort, D., Maslikowski, B., Warner, N., Kanno, N., Kim, H., Wang, Z., Moran, M. F., Oshima, R. G., Cardiff, R. D., and Muller, W. J. Grb2 and Shc adapter proteins play distinct roles in Neu (ErbB-2)-induced mammary tumorigenesis: implications for human breast cancer. Mol. Cell. Biol., *21:* 1540–1551, 2001.
25. Shen, K., and Novak, R. F. DDT stimulates c-erbB2, c-met, and STATS tyrosine phosphorylation, Grb2-Sos association, MAPK phosphorylation, and proliferation of human breast epithelial cells. Biochem. Biophys. Res. Commun., *231:* 17–21, 1997.
26. Tari, A. M., and Lopez-Berestein, G. GRB2: a pivotal protein in signal transduction. Semin. Oncol., *28:* 142–147, 2001.
27. Houen, G. Mammalian Cu-containing amine oxidases (CAOs): new methods of analysis, structural relationships, and possible functions. APMIS Suppl., *96:* 1–46, 1999.
28. Hirst, J., Bright, N. A., Rous, B., and Robinson, M. S. Characterization of a fourth adapter-related protein complex. Mol. Biol. Cell, *10:* 2787–2802, 1999.

Advances in Brief

Identification of Differentially Expressed Genes in Esophageal Squamous Cell Carcinoma (ESCC) by cDNA Expression Array: Overexpression of *Fra-1, Neogenin, Id-1,* and *CDC25B* Genes in ESCC[1]

Ying Chuan Hu, King Yin Lam, Simon Law, John Wong, and Gopesh Srivastava[2]

Departments of Pathology [Y. C. H., K. Y. L., G. S.] and Surgery [S. L., J. W.], Faculty of Medicine, The University of Hong Kong, Hong Kong

Abstract

Purpose: **This study aims to identify differentially expressed genes in esophageal squamous cell carcinoma (ESCC) through the use of a membrane-based cDNA array.**

Experimental Design: **Two newly established human ESCC cell lines (HKESC-1 and HKESC-2) and one corresponding to a morphologically normal, esophageal epithelium tissue specimen, prospectively collected from the HKESC-2-related patient, were screened in parallel using a cDNA expression array containing gene-specific fragments for 588 human genes spotted onto nylon membranes.**

Results: **The results of cDNA expression array showed that 53 genes were up-regulated 2-fold or higher and 8 genes were down-regulated 2-fold or higher in both ESCC cell lines at the mRNA level. Semiquantitative RT-PCR analysis of a subset of these differentially expressed genes gave results consistent with cDNA array findings. Four of the differentially expressed genes that belong to the categories of oncogenes/tumor suppressor genes (*Fra-1* and *Neogenin*) and cell cycle-related genes (*Id-1* and *CDC25B*) were studied more extensively for their protein expression by immunohistochemistry. The two ESCC cell lines and their corresponding primary tissues, 61 primary ESCC resected specimens and 16 matching, morphologically normal, esophageal epithelium tissues were analyzed. The immunostaining results showed that *Fra-1, Neogenin, Id-1,* and *CDC25B* were overexpressed in both ESCC cell lines and their corresponding primary tumors at the protein level, validating the microarray findings. The results of the clinical specimens showed that the *Fra-1* gene was overexpressed in ESCC compared with normal esophageal epithelium in 53 of 61 cases (87%), *Neogenin* in 57 of 61 cases (93%), *Id-1* in 57 of 61 cases (93%), and *CDC25B* in 48 of 61 cases (79%). Furthermore, the expression of *Fra-1, Neogenin,* and *Id-1* in ESCC correlated with tumor differentiation.**

Conclusions: **Overall, this study demonstrates that multiple genes are differentially expressed in ESCC and provides the first evidence that oncogenes *Fra-1* and *Neogenin* and cell cycle-related genes *Id-1* and *CDC25B* are overexpressed in ESCC.**

Introduction

Esophageal carcinoma is the ninth most common human cancer in the world and the second most common cancer in China (1). In Hong Kong, ESCC[3] accounts for ~90% of esophageal malignant tumors and is the sixth most common cause of cancer death (2). Despite advances in multimodality therapy, the overall 5-year survival rates for ESCCs still remain poor (3). The development of new treatment modalities, diagnostic technologies, and preventive approaches will require a better understanding of the molecular mechanisms underlying esophageal carcinogenesis. Although recent reports have documented alterations of a few oncogenes and tumor suppressor genes in ESCC, the molecular and genetic basis of esophageal carcinogenesis still remains largely unknown (4, 5).

With the emerging technology of cDNA array hybridization, it is now possible to screen for alterations in the expression of many genes simultaneously (6–8). Because the development and progression of cancer are accompanied by complex changes in patterns of gene expression (9, 10), the cDNA array technology provides a very useful tool for studying these complex processes (6). In this study, we used cDNA expression array hybridization to examine the expression of 588 genes in two newly established ESCC cell lines (HKESC-1 and HKESC-2) and one corresponding, morphologically normal, esophageal epithelium tissue specimen collected prospectively from the HKESC-2-related patient. By comparing gene expression levels between normal esophageal epithelium and the ESCC cell lines, we were able to identify the differentially expressed transcripts in ESCC. Subsequent semiquantitative RT-PCR analyses vali-

Received 3/20/01; revised 5/7/01; accepted 5/8/01.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Presented in part at the 91st Annual Meeting of the American Association for Cancer Research, Inc. San Francisco, CA, 2000.

[2] To whom requests for reprints should be addressed, at Department of Pathology, University of Hong Kong, University Pathology Building, Queen Mary Hospital, Pok Fu Lam Road, Hong Kong. Phone: (852) 2855-4859; Fax: (852) 2872-5197; E-mail: gopesh@pathology.hku.hk.

[3] The abbreviations used are: ESCC, esophageal squamous cell carcinoma; *Fra-1,* fos-related antigen 1; *CDC25B,* cell division cycle 25B; *Id-1,* inhibitor of differentiation 1 (inhibitor of DNA binding 1); IH, immunohistochemistry; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; RT-PCR, reverse transcription-PCR; *FPR1,* formyl peptide receptor 1; *RANTES,* regulated on activation, normal T expressed, and secreted; AP-1, activator protein-1; CDK, cyclin-dependent kinase; DAB, 3,3′-diaminobenzidine.

dated the cDNA array results. Expression of the protein products of four of these differentially expressed genes that belong to the categories of oncogenes/tumor suppressor genes (*Fra-1* and *Neogenin*) and cell cycle-related genes (*Id-1* and *CDC25B*) was further evaluated by IH in a large series of ESCC tumor specimens.

Materials and Methods

Cell Culture and Tissue Specimen. Two human ESCC cell lines, HKESC-1 and HKESC-2, have been established recently in our laboratory. HKESC-1 has been reported previously (11). Both cell lines were from Hong Kong Chinese patients with moderately differentiated ESCC; HKESC-1 was from a 47-year-old man, whereas HKESC-2 was from a 46-year-old woman. The squamous epithelial nature of HKESC-2 was confirmed by both electron microscopy (presence of tonofilaments and desmosomes) and immunohistochemical staining (positive for cytokeratins; data not shown). Both cell lines grew as adherent monolayers (11). HKESC-2 was maintained in the same conditions as HKESC-1 (11). Cells were harvested from passage 31 of HKESC-1 and passage 4 of HKESC-2 at 80–90% confluency, respectively.

One morphologically normal, esophageal epithelium tissue specimen, collected prospectively from the HKESC-2-related patient, was used as a control for the array experiment. For obtaining high-purity normal esophageal epithelium tissue specimen, the morphologically normal esophageal epithelium at least 5 cm away from the tumor margin was carefully dissected out from other tissues of the freshly resected esophagectomy specimen from the HKESC-2-related patient and evaluated microscopically. Unfortunately, the collected normal esophageal epithelium tissue from the HKESC-1-related patient could not be used as a control because the specimen was too small and only a small amount of RNA could be extracted from it.

cDNA Arrays, Probes, Hybridization, and Data Analysis. Atlas Human cDNA Expression Array membranes used in this study were purchased from Clontech (Palo Alto, CA). The membrane contained 10 ng of each gene-specific cDNA from 588 known genes and 9 housekeeping gene fragments (Fig. 1). Several plasmid and bacteriophage DNAs and blank spots are also included as negative and blank controls to confirm hybridization specificity.[4]

Total RNA was extracted using the Trizol reagent protocol (Life Technologies, Inc., Gaithersburg, MD) from the two ESCC cell lines (HKESC-1 and HKESC-2) and one corresponding, morphologically normal esophageal epithelium collected prospectively from the HKESC-2-related patient. mRNA was then isolated from the total RNA using the Straight A's mRNA Isolation System (Novagen, Madison, WI). The ^{32}P-labeled cDNA probes were generated by reverse transcription of 1 μg of mRNA of each sample in the presence of [α-^{32}P]dATP. Equal amounts of cDNA probes (3×10^6 cpm/μl) from the ESCC cell lines and normal esophageal epithelium were then hybridized to separate Atlas Human cDNA array membranes for 24 h at 42°C and washed according to the supplier's instructions. The array membranes were then exposed to X-ray film at −70°C for 2–5 days. Autoradiographic intensity was analyzed using AtlasImage analysis software (version 1.01; Clontech). The signal intensities were normalized by comparing the expression of housekeeping genes *GAPDH* (G12) and *HLA-C* (G14):

$$\text{Intensity ratio} = \frac{\text{Adjusted intensity on array} - \text{HKESC-1 or HKESC} - 2}{\text{Adjusted intensity on array} - \text{normal}} \times \text{Normalization coefficient}$$

$$\text{Adjusted intensity} = \text{Intensity} - \text{Background}$$

$$\text{Normalization coefficient} = \left[\left(\frac{\text{Adjusted intensity } \textit{GAPDH} \text{ on array} - \text{normal}}{\text{Adjusted intensity } \textit{GAPDH} \text{ on array} - \text{HKESC-1 or HKESC-2}}\right)\left(\frac{\text{Adjusted intensity HLA} - \text{Con array} - \text{normal}}{\text{Adjusted intensity HLA} - \text{Con array} - \text{HKESC-1 or HKESC-2}}\right)\right] \div 2$$

Genes were considered to be up-regulated when the intensity ratio between expression in the ESCC cell lines compared with normal esophageal epithelium was 2-fold or greater. Genes were labeled as down-regulated when the ratio between normal and ESCC cell lines was 2-fold or higher.

RT-PCR. cDNA was generated using 1 μg of total RNA from the two ESCC cell lines (HKESC-1 and HKESC-2) and one corresponding, morphologically normal, esophageal epithelium collected prospectively from the HKESC-2-related patient as template and 2.5 mM Oligo $d(T)_{16}$ primers in a 20-μl reaction mixture, and the reverse transcription was carried out at 42°C for 1 h, followed by 95°C for 10 min using the GeneAmp RNA PCR Core kit (Perkin-Elmer, Branchburg, NJ). Two μl of cDNA were amplified in a 25-μl PCR reaction mixture containing 1× PCR buffer, 1.9 or 2.4 or 2.9 mM $MgCl_2$, 0.5 μM primers, 0.18 mM deoxynucleotides triphosphates, 1 unit of AmpliTaq Gold DNA polymerase with hot-start PCR as follows: 95°C for 10 min, followed by 25–35 cycles of 1 min denaturation at 94°C, 1 min annealing at 60°C (for primers of *cyclin D1, Id-1, CDC25B, FPR1, RANTES,* and *GAPDH*) or 62°C (for primers of *Fra-1*) or 65°C (for primers of *Neogenin*), 1 min extension at 72°C. Finally, PCR products were fully extended by incubating at 72°C for 10 min. The PCR reagents were purchased from Perkin-Elmer.

The sequences of gene specific primers for RT-PCR were the same as those of cDNA array (data not shown because of the copyright agreement by Clontech, Palo Alto, CA), except for the primers specific for *Fra-1,* which were the same as described before (12). All of the primers were synthesized by Integrated DNA Technologies, Inc. (Coralville, IA). The cycle number was optimized for each gene-specific primer pair to ensure that amplification was in the linear range and the results were semiquantitative. Twelve μl of PCR product were visualized by

[4] A complete list of the 588 genes with array positions and GenBank accession numbers of the Atlas Human Expression Array used here can also be accessed through the web site http://www.clontech.com/clontech/APR97UPD/Atlaslist.html.

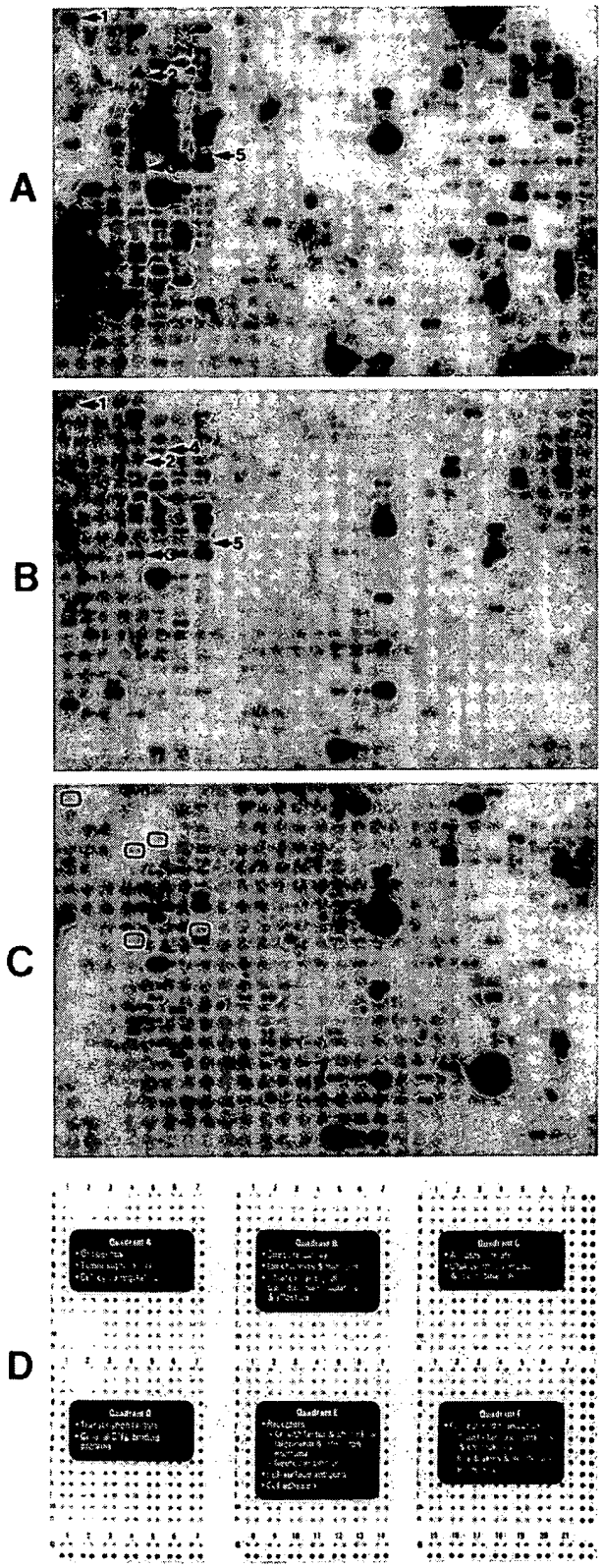


Fig. 1 *A–C*, gene expression profiles of two human ESCC cell lines HKESC-1 (*A*) and HKESC-2 (*B*) and one morphologically normal esophageal epithelium (*C*) from the HKESC-2-related patient using the Atlas Human cDNA Expression Array. Some of the differentially expressed genes are indicated: *1*, c-*myc* (A1a); *2*, *Fra-1* (A4f); *3*, *Neogenin* (A4n); *4*, *Id-1* (A5e); and *5*, *CDC25B* (A7m). *D*, schematic diagram of Atlas Human cDNA Expression Array. The array contains 588 human genes spotted in duplicate and divided into six functional categories (*Quadrants A–F*). Three blank (G1, G8, and G15) and nine negative (G2–4, G9–11, and G16–18) controls are included to confirm hybridization specificity. Nine housekeeping genes (G5–7, G12–14, and G19–21) are also included in the array for normalizing mRNA abundance. Genomic DNA spots (*dark dots*) serve as orientation marks to facilitate in the determination of the coordinates of hybridization signals. A complete gene list with array coordinates and GenBank accession numbers is available at world wide web site http://www.clontech.com/clontech/APR97 UPD/Atlaslist.html.

electrophoresis on a 2% agarose gel stained with ethidium bromide and quantitated by densitometry using a dual-intensity transilluminator equipped with Gelworks 1D Intermediate software (version 2.51).

Collection of Tissues and Clinicopathological Data. The tissues were obtained from 61 (50 men and 11 women) patients with ESCC resected between 1996 and 1998 in Queen Mary Hospital, The University of Hong Kong. The patients' ages ranged from 41 to 83 years, with a mean age of 65 years. The specimens were dissected and examined in the fresh state. Representative tissue specimens from tumors and matching normal esophageal epithelium tissues were snap-frozen in liquid nitrogen and stored at −80°C. Other representative blocks were taken and processed in paraffin for histological examination. The carcinomas were found in the upper (n = 10; 16%), middle (n = 35; 57%), and lower (n = 16; 26%) third of the esophagus. The median length of the tumors was 5.5 cm (range, 1–11). The histology of the carcinomas was reviewed according to the criteria described previously (13). The ESCC tumors were well differentiated in 20 (33%) cases, moderately differentiated in 29 (48%), and poorly differentiated in 12 (20%). The carcinomas were staged according to the Tumor-Node-Metastasis classification (14). Many tumors were stage III (n = 35, 57%) or II (n = 23, 38%); of the remainder, one was stage I and two were stage IV.

Immunohistochemistry. Expression of *Fra-1*, *Neogenin*, *Id-1*, and *CDC25B* was investigated by the streptavidin-biotin-peroxidase complex method. Briefly, 6-μm frozen sections were cut from two pellets harvested from cultured cell lines HKESC-1 and HKESC-2, the cell lines' corresponding primary tissues, 61 primary ESCC tumors, and 16 matching, morphologically normal, esophageal epithelium specimens. After endogenous peroxidase activity was quenched and nonspecific binding was blocked, polyclonal rabbit anti-Fra-1, goat anti-Neogenin, rabbit anti-Id-1, and goat anti-CDC25B antibodies (Santa Cruz Biotechnology, Santa Cruz, CA) were incubated at 4°C overnight at a dilution of 1:100 for *Fra-1*, 1:40 for *Neogenin* and *Id-1*, and 1:30 for *CDC25B*, respectively. The secondary antibody was biotinylated swine anti-rabbit (for *Fra-1* and *Id-1*) or rabbit anti-goat (for *Neogenin* and *CDC25B*) antibody (DAKO, Glostrup, Denmark) used at a dilution of 1:200 for 30 min at 37°C. After washing, sections were incubated with StreptABComplex/horseradish peroxidase (DAKO; 1:100 dilution) for 30 min at 37°C. Negative controls were performed by replacing the primary antibody by normal serum.

Each section was independently assessed by two histopathologists without prior knowledge of the patients' other data. Scoring was based on the percentage of positive cells. The staining was identified as: −, no expression; +, <10% of cells were stained; ++, 10–50% of cells stained; and +++, >50% of cells stained. From ++ to +++ was defined as overexpression.

Statistical Analysis. Comparisons between groups were performed using the χ^2 test and t test when appropriate. $P < 0.05$ was used to determine statistical significance. All statistical tests were performed with the GraphPad Prism software version 3.0 (GraphPad Software, Inc., San Diego, CA).

Results and Discussion

In this study, we first used cDNA expression array hybridization to identify genes that were differentially expressed in ESCC compared with normal esophageal epithelium. Two newly established ESCC cell lines in our laboratory were selected for cDNA array analysis to assure large quantities of high-purity tumor mRNA. The comparison of the autoradiographic intensities between ESCC cell lines and normal esophageal epithelium showed that 65 and 59 genes were up-regulated 2-fold or higher and 11 and 21 genes were down-regulated 2-fold or more in HKESC-1 and HKESC-2, respectively. Among these differentially expressed genes, 53 genes were up-regulated and 8 genes were down-regulated in both cell lines (Fig. 1 and Table 1). No signals were visible in the three blank spots (G1, G8, and G15) and nine negative control spots (G2–4, G9–11, and G16–18; Fig. 1), indicating that the cDNA array hybridization was highly specific. Among the 61 differentially expressed genes in both cell lines, 49 genes such as *Fra-1, Neogenin, Id-1,* and *CDC25B* genes were identified as differentially expressed in ESCC for the first time; 12 other differentially expressed genes have been described to be overexpressed in ESCC previously. The genes overexpressed in both of the ESCC cell lines belong to the categories of oncogenes/tumor suppressor genes, cell cycle-related genes, genes for DNA synthesis, DNA binding genes, or apoptosis-related genes (Table 1). The 8 genes that were down-regulated in both ESCC cell lines comprised genes for signal transduction (*guanine nucleotide regulatory protein NET1, protein kinase C-β II, cAMP-dependent protein kinase catalytic d-subunit and EPLG3*), genes for signaling proteins (*RANTES protein T-cell specific, Somatomedin A,* and *FPR1*), and the gene for MAL protein. These findings demonstrated that multiple genes are differentially expressed in ESCC at mRNA level.

To further validate the cDNA array approach, we performed semiquantitative RT-PCR to analyze the expression levels of 8 genes, *cyclin D1, Fra-1, Neogenin, Id-1, CDC25B, FPR1, RANTES,* and *GAPDH*. The results of RT-PCR analysis (Fig. 2) were consistent with the expression profiles obtained through cDNA array hybridization (Fig. 1).

Genes that belong to the categories of oncogenes/tumor suppressor genes and cell cycle-related genes are often implicated in the pathogenesis of various cancers (4, 5, 15). Significantly, a number of the differentially expressed genes identified by cDNA array hybridization in both ESCC cell lines belong to these categories. Four of these differentially expressed genes that were identified for the first time to be overexpressed in ESCC in this study, the oncogenes *Fra-1* and *Neogenin* and the cell cycle related genes *Id-1* and *CDC25B,* were selected for more detailed study for their protein expression in a large series of ESCC tumor specimens by IH. Moreover, these genes have been reported to be overexpressed in other tumor cell lines or primary tumors (16–24). The other consideration for selecting these particular genes for more extensive study was that the suitable antibodies of these genes were commercially available for the IH studies. The protein expression of these four genes was investigated in the two ESCC cell lines and their corresponding primary tissues, 61 primary ESCC tumors, and 16 matching, morphologically normal, esophageal epithelium specimens. The results of immunostaining are summarized in Tables 2 and 3 and are shown in Fig. 3. The protein products of *Fra-1, Neogenin, Id-1,* and *CDC25B* genes were found to be overexpressed in both the ESCC cell lines and their corresponding primary tumors (Table 2), validating the cDNA array results.

Fra-1 is one component of the AP-1 complex (25). The AP-1 components are considered to play key roles in signal transduction pathways involved in complex cellular growth, differentiation, and tumorigenesis (16). Previous studies indicated that increased AP-1 activity is a necessary event in the transformation of mouse epidermal cells (26, 27). *Fra-1* overexpression has been found in kidney and thyroid cancer (16, 17). These observations suggest that *Fra-1* overexpression might play an important role in malignant transformation of epithelial cells. In the present work, *Fra-1* mRNA overexpression was detected in both ESCC cell lines by cDNA array analysis (Fig. 1) and RT-PCR (Fig. 2*B*). Also, the majority of ESCC tumors (53 of 61, 87%; Table 3; Fig. 3*B*) had enhanced expression of *Fra-1*. Fra-1 protein expression was localized in the nuclei of ESCC tumor cells (Fig. 3*B*). In contrast, morphologically normal, esophageal epithelium tissues showed low expression of Fra-1. The expression of Fra-1 was often focal in morphologically normal esophageal epithelium and always restricted to the basal cell layer (Fig. 3*A*). The well or moderately differentiated ESCC showed more intense expression of Fra-1 than poorly differentiated ones ($P < 0.0001$; Table 3).

Neogenin encodes a 1461-amino acid protein with 50% amino acid identity to *DCC* (*deleted in colorectal cancer*; Ref. 18). It has been suggested to play an integral role in regulating differentiation and/or cell migration events within many embryonic and adult tissues (28). *Neogenin* expression has been detected at low levels in many adult tissues but not including esophagus (18). Overexpression of *Neogenin* has been observed in a wide variety of human cancer cell lines from cancers of breast, pancreas, brain, cervix, colon, and rectum (18). However, there is no information about the status of *Neogenin* expression in human primary cancers including esophageal cancer. In the current study, our cDNA array (Fig. 1) and RT-PCR (Fig. 2*B*) results showed that *Neogenin* mRNA was overexpressed in both ESCC cell lines. Neogenin protein overexpression was noted in 93% (57 of 61) of ESCCs (Table 3; Fig. 3*D*). The expression was localized in the cytoplasm of tumor cells (Fig. 3*D*). In contrast, the expression of Neogenin protein in morphologically normal esophageal epithelium was negative or negligible and was restricted to the highly proliferative basal cells (Fig. 3*C*). The well or moderately differentiated ESCC

Table 1 List of differentially expressed genes in the ESCC cell lines HKESC-1 and HKESC-2 when compared with one corresponding, morphologically normal, esophageal epithelium tissue specimen (N) from the HKESC-2-related patient using cDNA expression array

Location	Name of gene	Intensity Ratio HKESC-1/N	Intensity Ratio HKESC-2/N
Genes up-regulated in both ESCC cell lines			
Oncogene/Tumor suppressor genes			
A1a	*c-myc*	2.5	3.0
A2b	*IGFBP-2*	6.9	6.0
A3b	*Snon*	7.2	6.0
A3i	*rhoA* (MDR protein)	2.8	3.9
A3k	*DCC*	5.1	3
A4b	*APC*	2.6	3.9
A4c	*BRCA2*	30029/0	42730/0
A4f	*Fra-1*	10.8	7.8
A4g	*Ezrin*	21936/0	36475/0
A4h	*JUN-D*	35.4	60.6
A4j	*PEP1*	3.6	4.8
A4k	*EB1*	2.6	3.0
A4l	*C-CBL*	2.7	2.2
A4m	*Smad1*	3.2	3.0
A4n	*Neogenin*	2.4	3.2
Cell cycle-related genes			
A5e	*Id-1*	3.4	3.7
A5g	*P58/GTA1*	4.0	2.2
A6g	*Cyclin D1*	2.1	2.3
A6l	*Cyclin B1*	2.4	3.2
A6m	*Cyclin E*	3.5	3.3
A7b	*Cyclin G2*	2.4	3.1
A7d	*p35*	6.9	7.2
A7l	*C-1*	2.6	2.6
A7m	*CDC25B*	2.0	2.5
Apoptosis-associated genes			
C1i	*Adenosine A1 receptor*	22.2	5.9
C4l	*Apopain*	8.7	13.7
Genes for DNA synthesis/repair/recombination proteins			
C6d	*XRCC1*	3.2	2.9
C6l	*DNA Topoisomerase II*	14759/0	22201/0
C7n	*Dnase X*	5.1	4.4
Genes for DNA binding/transcription factors			
D1c	*CCAT-binding protein*	9.9	8.8
D1d	*Id-3*	6.5	6.6
D1e	*BTEB2*	7.4	9.0
D1g	*Id-2*	9.8	7.6
D1l	*TAX*	3.0	4.4
D1n	*CNBP*	4.3	2.1
D2a	*CCAAT displacement protein*	4.2	4.5
D2c	*APRF*	33987/0	29011/0
D2d	*hSNF2b*	42.2	40.7
D2f	*TAXREB67*	39440/0	49383/0
D2i	*TCF5*	37478/0	23080/0
D3a	*hSNF2a*	7.1	4.2
D3b	*DB1*	316.5	225.0
D3c	*D-binding protein*	37796/0	21621/0
D3g	*PAX-8*	13.6	7.2
D3j	*P15 subunit*	38725/0	40864/0
D3k	*Guanine nucleotide-binding protein G-S*	33.0	49.2
D4c	*AP-2*	7.1	4.9
D4j	*NF-E1*	34506/0	15924/0
D5k	*PAX3*	24.8	10.6
D7k	*TAFII31*	7.0	4.8
Genes for signal proteins			
F5a	*NGF-2*	16.7	4.6
F5b	*MIP2α*	39223/0	19827/0
F5f	*IL-8*	39239/0	11638/0
Genes down-regulated in both ESCC cell lines			
Gene for iron channel/transport protein			
B1b	*MAL protein*	1/3.3	0/11989
Genes for signal transduction			
B4g	*Guanine nucleotide regulatory protein NET1*	0/28564	1/23.7
B5j	*Protein kinase c-β II*	0/29588	0/29588
B6b	*cAMP-dependent protein kinase α-subunit*	0/41060	1/304
B6n	*EPLG3*	0/22304	1/2.5
Genes for cell signaling proteins			
F1a	*Somatomedin A*	0/19560	0/19560
F1k	*FMLP-related receptor I*	0/29220	0/29220
F2j	*RANTES protein T-cell specific*	0/36868	1/3.2

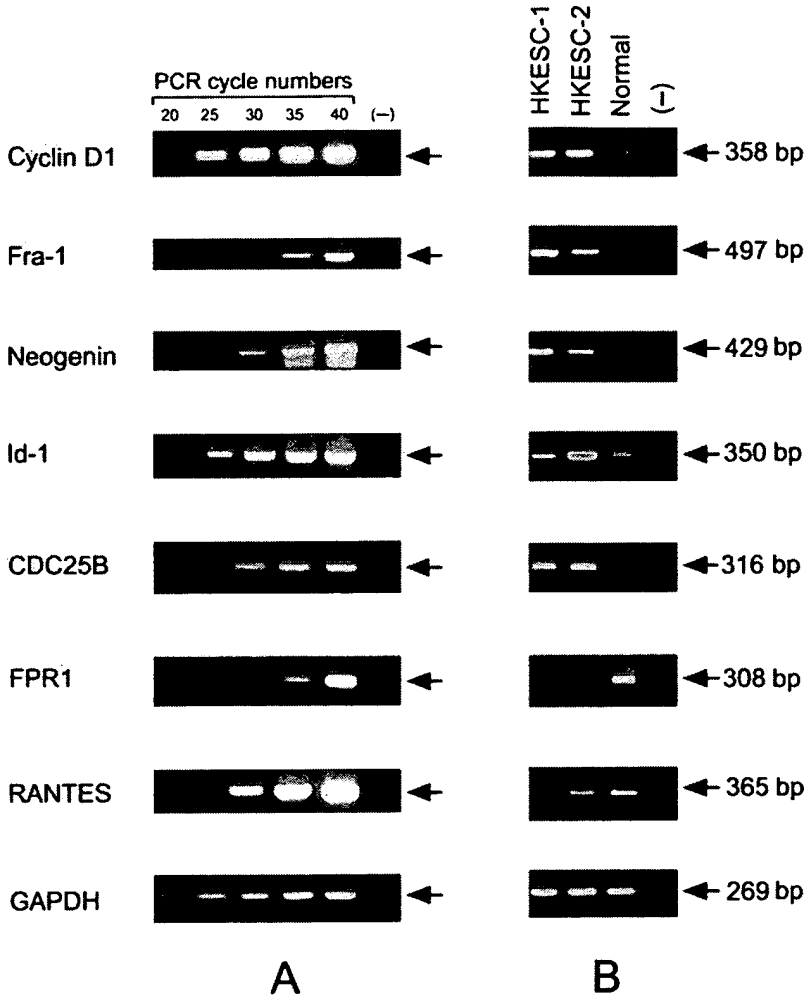


Fig. 2 RT-PCR analysis of *cyclin D1, Fra-1, Neogenin, Id-1, CDC25B, FPR1, RANTES,* and *GAPDH* genes in ESCC cell lines HKESC-1 and HKESC-2 and one corresponding, morphologically normal epithelium (*Normal*) from the HKESC-2-related patient. *A*, determination of optimal number of PCR cycles for different gene-specific primer pairs. mRNA from HKESC-1 was used to determine the optimal number of PCR cycles for genes *cyclin D1, Fra-1, Neogenin, Id-1, CDC25B,* and *GAPDH*. mRNA from the normal esophageal epithelium was used to determine the optimal number of PCR cycles for genes *FPR1* and *RANTES*. *B*, expression of *cyclin D1* (25 cycles), *Fra-1* (32 cycles), *Neogenin* (30 cycles), *Id-1* (25 cycles), *CDC25B* (28 cycles), *FPR1* (35 cycles), *RANTES* (28 cycles), and *GAPDH* (25 cycles) genes in two ESCC cell lines HKESC-1 and HKESC-2 and one corresponding, morphologically normal esophageal epithelium (*Normal*) from the HKESC-2-related patient.

Table 2 Summary of immunohistochemical staining results in ESCC cell lines and their corresponding primary tissue specimens

	Cell lines		Primary tissues[a]			
	HKESC-1	HKESC-2	T_1	N_1	T_2	N_2
Fra-1	+++[b]	+++	+++	+	++	+
Neogenin	+++	+++	+++	–	++	+
Id-1	+++	+++	+++	++	++	–
CDC25B	+++	+++	+++	–	++	+

[a] T, ESCC tumor; N, morphologically normal esophageal epithelium.
[b] Expression: –, no expression; +, <10% cells positive; ++, ≥10% and <50% cells positive; +++, ≥50% cells positive. ++ to +++ was considered as overexpression.

showed more intense expression of *Neogenin* than poorly differentiated ones ($P = 0.0047$; Table 3).

Id-1 is a cell cycle-related gene that encodes a helix-loop-helix protein. *Id-1* plays an important role not only in suppressing cellular differentiation but also in enhancing cellular proliferation (29, 30). Generally, Id-1 protein is highly expressed in growing cells, and its expression is down-regulated upon differentiation in many cell types. Although Id-1 is expressed in a variety of fetal tissues and overexpressed in tumors from brain and lung (19), its expression in ESCC is unknown. In this study, we observed that *Id-1* mRNA was overexpressed in both ESCC cell lines (Figs. 1 and 2*B*). Also, Id-1 protein overexpression was frequent in human primary ESCC tumors (57 of 61, 93%; Table 3; Fig. 3*F*). The Id-1 protein was localized in the cytoplasm of tumor cells (Fig. 3*F*). In contrast, the expression of Id-1 protein in morphologically normal esophageal epithelium was either negative or negligible and was restricted to the basal and parabasal cells (Fig. 3*E*). The well or moderately differentiated ESCC showed more intense expression of *Id-1* than poorly differentiated ones ($P = 0.0156$; Table 3).

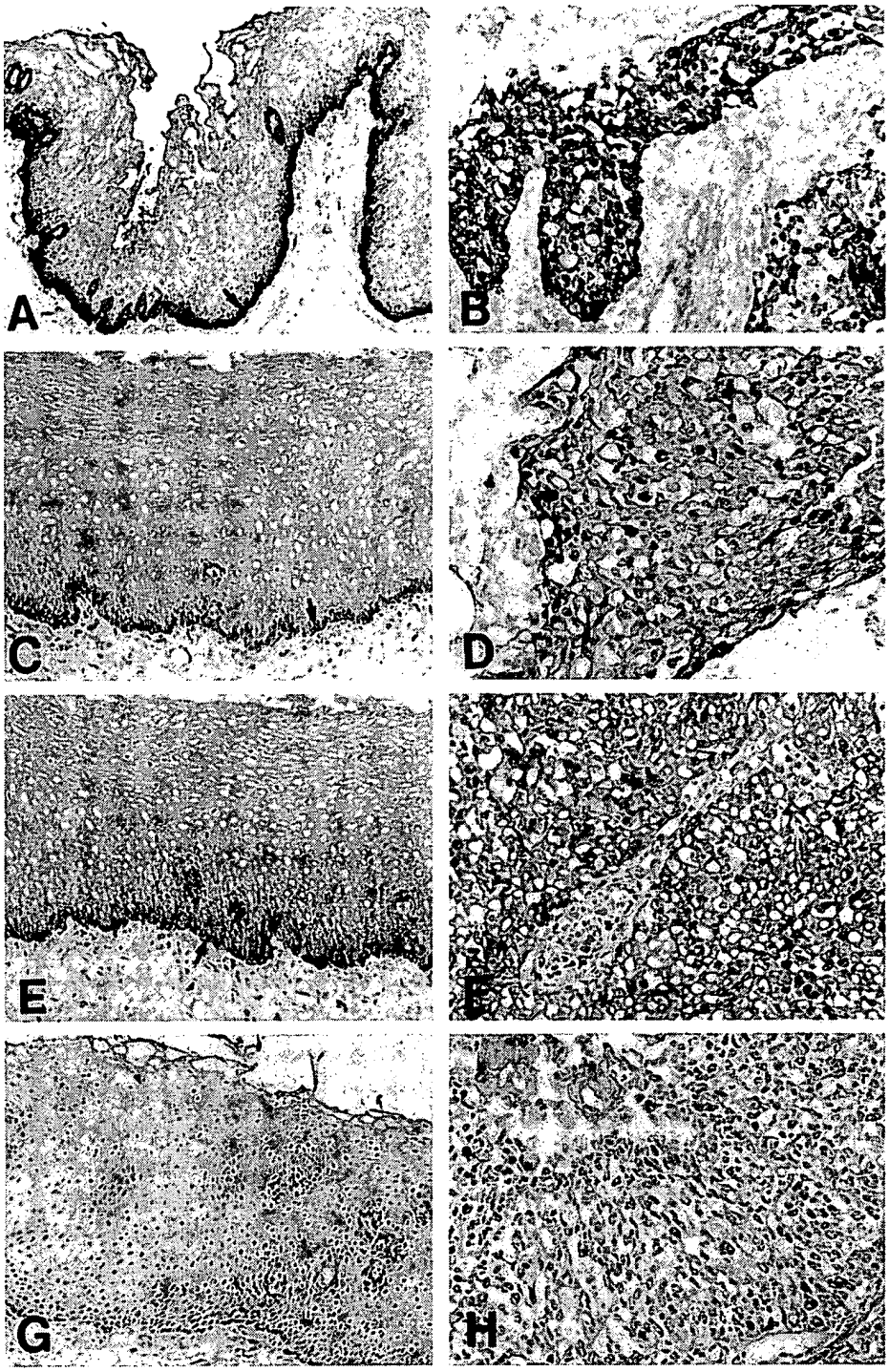


Fig. 3 Photomicrographs of *Fra-1*, *neogenin*, *Id-1*, and *CDC25B* expression by IH in morphologically normal esophageal epithelium and ESCC. *A*, *Fra-1* IH in morphologically normal esophageal epithelium showing that *Fra-1* expression was restricted to the basal cell layer (*arrow*); DAB ×160. *B*, *Fra-1* IH in ESCC showing that the nuclei of tumor cells are strongly positive for *Fra-1*; DAB × 400. *C*, *Neogenin* IH in morphologically normal esophageal epithelium showing that *Neogenin* expression was restricted to the basal cell layer (*arrow*); DAB ×330. *D*, *Neogenin* IH in ESCC showing *Neogenin* expression in the cytoplasm of tumor cells; DAB ×500. *E*, *Id-1* IH in morphologically normal esophageal epithelium showing *Id-1* expression was restricted to the basal (*arrow*) and parabasal (*arrowhead*) cell layers; DAB ×330. *F*, *Id-1* IH in ESCC showing Id-1 expression in the cytoplasm of tumor cells; DAB ×500. *G*, *CDC25B* IH in morphologically normal esophageal epithelium showing lack of immunoreactivity; DAB × 330. *H*, *CDC25B* IH in ESCC showing *CDC25B* expression in the nuclei of most tumor cells; DAB ×500.

CDC25B is a cell cycle-related gene. Its product is a phosphatase that catalyzes the removal of inhibitory phosphate from the CDK family of proteins (31). *CDC25B* can dephosphorylate threonine 14, tyrosine 15, or both on CDKs and activate cyclin/CDK complexes to stimulate cell proliferation (32). *In vitro* transforming experiments have demonstrated that *CDC25B* is also a potential oncogene (20). Overexpression of *CDC25B* has been found in cancers arising from breast (20), stomach (21), lung (22), and head and neck (23), and in non-Hodgkin's lymphoma (24). In this study, we demonstrated that the mRNA of *CDC25B* was highly expressed in both ESCC cell lines by cDNA array (Fig. 1) and RT-PCR (Fig. 2*B*). Furthermore, *CDC25B* was overexpressed in 79% (48 of 61) primary ESCC tumors by IH (Table 3; Fig. 3*H*). CDC25B protein

Table 3 Summary of IH staining results in clinical ESCC tumors and normal esophageal epithelium tissues

Diagnosis	Fra-1				P	Neogenin				P	Id-1				P	CDC25B				P
	−[a]	+	++	+++		−	+	++	+++		−	+	++	+++		−	+	++	+++	
Normal (n = 16)	1	15	0	0	<0.0001	6	10	0	0	<0.0001	3	12	1	0	<0.0001	9	7	0	0	<0.0001
Carcinoma (n = 61)	4	4	22	31		3	1	7	50		3	1	8	49		10	3	19	29	
Well (n = 20)	0	0	9	11		1	0	2	17		0	1	3	16		2	2	7	9	
Moderate (n = 29)	1	0	11	17	<0.0001	0	0	1	28	0.0047	0	0	3	26	0.0156	6	1	10	12	0.5720
Poor (n = 12)	3	4	2	3		2	1	4	5		3	0	2	7		2	0	2	8	

[a] Expression: −, no expression; +, <10% cells positive; ++, ≥10% and <50% cells positive; +++, ≥50% cells positive. ++ to +++ was considered as overexpression.

expression was localized mainly in the nuclei of tumor cells (Fig. 3*H*). On the other hand, the expression of *CDC25B* in morphologically normal esophageal epithelium tissues was either negative or very weak (Fig. 3*G*). In the case of *CDC25B*, there was no correlation between gene expression and ESCC differentiation (P = 0.5720; Table 3).

In summary, all four of the genes selected for further study demonstrated a significantly higher incidence of overexpression in primary ESCCs than morphologically normal esophageal epithelium tissues ($P < 0.0001$; Table 3). Furthermore, three of them, *Fra-1, Neogenin,* and *Id-1* were more highly expressed in tumors with greater differentiation. *CDC25B* did not demonstrate this correlation (Table 3). The expression of these genes did not correlate to age at presentation or gender of patients or tumor site, size, or stage. The differentiation of squamous cell carcinoma bears no relationship with the stage of the tumor (4). In this study, the expression of *Fra-1, Neogenin,* and *Id-1* was more often noted in the well/moderately differentiated squamous cell carcinoma. This is consistent with the theory that poorly differentiated squamous cell carcinoma arises at the early stage of carcinogenesis. In the later stages of tumor progression, the squamous cell carcinoma becomes more mature in appearance (well/moderately differentiated).

Unfortunately, the complete follow-up data were available only for some of these patients. Nevertheless, all these patients died within 2 years of resection of the primary tumors. Also, the *Fra-1, Neogenin, Id-1,* and *CDC25B* were highly expressed in squamous cell carcinomas. Thus, it is unlikely that the expression of these genes acts as an independent prognostic factor in these tumors.

Overall, our data demonstrate that multiple genes are differentially expressed in ESCC and show for the first time that oncogenes *Fra-1* and *Neogenin* and cell cycle-related genes *Id-1* and *CDC25B* are overexpressed in ESCC. Additional studies are required to determine the roles of these and other differentially expressed genes in the molecular pathogenesis of ESCC.

References

1. Blot, W. J. Esophageal cancer trends and risk factors. Semin. Oncol., *21:* 403–410, 1994.

2. Lam, K. Y., and Ma, L. Pathology of esophageal cancers: local experience and current insights. Chin. Med. J., *10:* 459–464, 1997.

3. Law, S. Y., Fok, M., Cheng, S. W., and Wong, J. A comparison of outcome after resection for squamous cell carcinomas and adenocarcinomas of the esophagus and cardia. Surg. Gynecol. Obstet., *175:* 107–112, 1992.

4. Lam, K. Y., Tsao, S. W., Zhang, D., Law, S., He, D., Ma, L., and Wong, J. Prevalence and predictive value of p53 mutation in patients with esophageal squamous cell carcinomas: a prospective clinico-pathological study and survival analysis of 70 patients. Int. J. Cancer, *74:* 212–219, 1997.

5. Montesano, R., Hollstein, M., and Hainaut, P. Genetic alterations in esophageal cancer and their relevance to etiology and pathogenesis: a review. Int. J. Cancer, *69:* 225–235, 1996.

6. DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A., and Trent, J. M. Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat. Genet., *14:* 457–460, 1996.

7. Fuller, G. N., Rhee, C. H., Hess, K. R., Caskey, L. S., Wang, R., Bruner, J. M., Yung, W. K. A., and Zhang, W. Reactivation of insulin-like growth factor binding protein 2 expression in glioblastoma multiforme: a revelation by parallel gene expression profiling. Cancer Res., *59:* 4228–4232, 1999.

8. Shim, C., Zhang, W., Rhee, C. H., and Lee, J. H. Profiling of differentially expressed genes in human primary cervical cancer by complementary DNA expression array. Clin. Cancer Res., *4:* 3045–3050, 1998.

9. Vogelstein, B., and Kinzler, K. W. The multistep nature of cancer. Trends Genet., *9:* 138–141, 1993.

10. Zhang, L., Zhou, W., Velculescu, V. E., Kern, S. E., Hruban, R. H., Hamilton, S. R., Vogelstein, B., and Kinzler, K. W. Gene expression profiles in normal and cancer cells. Science (Wash. DC), *276:* 1268–1272, 1997.

11. Hu, Y. C., Lam, K. Y., Wan, T. S. K., Fang, W. G., Ma, E. S. K., Chan, L. C., and Srivastava, G. Establishment and characterization of HKESC-1, a new cancer cell line from human esophageal squamous cell carcinoma. Cancer Genet. Cytogenet., *118:* 112–120, 2000.

12. Risse-Hackl, G., Adamkiewicz, J., Wimmel, A., and Schuermann, M. Transition from SCLC to NSCLC phenotype is accompanied by an increased TRE-binding activity and recruitment of specific AP-1 proteins. Oncogene, *16:* 3057–3068, 1998.

13. Lam, K. Y., He, D., Ma, L., Zhang, D., Ngan, H. Y., Wan, T. S., and Tsao, S. W. Presence of human papillomavirus in esophageal squamous cell carcinomas of Hong Kong Chinese and its relationship with *p53* gene mutation. Hum. Pathol., *28:* 657–663, 1997.

14. Beahrs, O. H., Henson, D. E., Hutter, R. V. P., and Kennedy, B. J. (eds.). Manual for Staging of Cancer, Ed. 4, pp. 57–59. Philadelphia: J. B. Lippincott Co., 1992.

15. Hunter, T. Oncoprotein networks. Cell, *88:* 333–346, 1997.

16. Battista, S., de Nigris, F., Fedele, M., Chiappetta, G., Scala, S., Vallone, D., Pierantoni, G. M., Megar, T., Santoro, M., Viglietto, G., Verde, P., and Fusco, A. Increase in AP-1 activity is a general event in thyroid cell transformation *in vitro* and *in vivo*. Oncogene, *17:* 377–385, 1998.

17. Urakami, S., Tsuchiya, H., Orimoto, K., Kobayashi, T., Igawa, M., and Hino, O. Overexpression of members of the AP-1 transcriptional factor family from an early stage of renal carcinogenesis and inhibition of cell growth by *AP-1* gene antisense oligonucleotides in the *Tsc2* gene

mutant (Eker) rat model. Biochem. Biophys. Res. Commun., *241:* 24–30, 1997.

18. Meyerhardt, J. A., Look, A. T., Bigner, S. H., and Fearon, E. R. Identification and characterization of *Neogenin*, a DCC-related gene. Oncogene, *14:* 1129–1136, 1997.

19. Zhu, W., Dahmen, J., Bulfone, A., Rigolet, M., Hernandez, M. C., Kuo, W. L., Puelles, L., Rubenstein, J. L., and Israel, M. A. *Id* gene expression during development and molecular cloning of the human *Id-1* gene. Brain Res. Mol. Brain Res., *30:* 312–326, 1995.

20. Galaktionov, K., Lee, A. K., Eckstein, J., Draetta, G., Meckler, J., Loda, M., and Beach, D. CDC25 phosphatases as potential human oncogenes. Science (Wash. DC), *269:* 1575–1577, 1995.

21. Kudo, Y., Yasui, W., Ue, T., Yamamoto, S., Yokozaki, H., Nikai, H., and Tahara, E. Overexpression of cyclin-dependent kinase-activating CDC25B phosphatase in human gastric carcinomas. Jpn. J. Cancer Res., *88:* 947–952, 1997.

22. Wu, W., Fan, Y. H., Kemp, B. L., Walsh, G., and Mao, L. Overexpression of cdc25A and cdc25B is frequent in primary non-small cell lung cancer but is not associated with overexpression of c-myc. Cancer Res., *58:* 4082–4085, 1998.

23. Gasparotto, D., Maestro, R., Piccinin, S., Vukosavljevic, T., Barzan, L., Sulfaro, S., and Boiocchi, M. Overexpression of CDC25A and CDC25B in head and neck cancers. Cancer Res., *57:* 2366–2368, 1997.

24. Hernandez, S., Hernandez, L., Bea, S., Cazorla, M., Fernandez, P. L., Nadal, A., Muntane, J., Mallofre, C., Montserrat, E., Cardesa, A., and Campo, E. cdc25 cell cycle-activating phosphatases and c-myc expression in human non-Hodgkin's lymphomas. Cancer Res., *58:* 1762–1767, 1998.

25. Cohen, D. R., and Curran, T. *Fra-1:* a serum-inducible, cellular immediate-early gene that encodes a fos-related antigen. Mol. Cell. Biol., *8:* 2063–2069, 1988.

26. Li, J. J., Dong, Z., Dawson, M. I., and Colburn, N. H. Inhibition of tumor promoter-induced transformation by retinoids that transrepress AP-1 without transactivating retinoic acid response element. Cancer Res., *56:* 483–489, 1996.

27. Joseloff, E., and Bowden, G. T. Regulation of the transcription factor AP-1 in benign and malignant mouse keratinocyte cells. Mol. Carcinog., *18:* 26–36, 1997.

28. Keeling, S. L., Gad, J. M., and Cooper, H. M. Mouse neogenin, a DCC-like molecule, has four splice variants and is expressed widely in the adult mouse and during embryogenesis. Oncogene, *15:* 691–700, 1997.

29. Norton, J. D., Deed, R. W., Craggs, G., and Sablitzky, F. Id helix-loop-helix proteins in cell growth and differentiation. Trends Cell Biol., *8:* 58–65, 1998.

30. Desprez, P. Y., Hara, E., Bissell, M. J., and Campisi, J. Suppression of mammary epithelial cell differentiation by the helix-loop-helix protein Id-1. Mol. Cell. Biol., *15:* 3398–3404, 1995.

31. Gautier, J., Solomon, M. J., Booher, R. N., Bazan, J. F., and Kirschner, M. W. cdc25 is a specific tyrosine phosphatase that directly activates p34cdc2. Cell, *67:* 197–211, 1991.

32. Galaktionov, K., and Beach, D. Specific activation of cdc25 tyrosine phosphatases by B-type cyclins: evidence for multiple roles of mitotic cyclins. Cell, *67:* 1181–1194, 1991.

1: Cancer Res. 1995 Mar 1;55(5):1168-75.

Malignant transformation of the human endometrium is associated with overexpression of lactoferrin messenger RNA and protein.

Walmer DK, Padin CJ, Wrona MA, Healy BE, Bentley RC, Tsao MS, Kohler MF, McLachlan JA, Gray KD.

Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, North Carolina 27710.

In the mouse uterus, lactoferrin is a major estrogen-inducible uterine secretory protein, and its expression correlates directly with the period of peak epithelial cell proliferation. In this study, we examine the expression of lactoferrin mRNA and protein in human endometrium, endometrial hyperplasias, and adenocarcinomas using immunohistochemistry, Western immunoblotting, and Northern and in situ RNA hybridization techniques. Our results reveal that lactoferrin is expressed in normal cycling endometrium by a restricted number of glandular epithelial cells located deep in the zona basalis. Two thirds (8 of 12) of the endometrial adenocarcinomas examined overexpress lactoferrin. This tumor-associated increase in lactoferrin expression includes an elevation in the mRNA and protein of individual cells and an increase in the number of cells expressing the protein. In comparison, only 1 of the 10 endometrial hyperplasia specimens examined demonstrates an increase in lactoferrin. We also observe distinct cytoplasmic and nuclear immunostaining patterns under different fixation conditions in both normal and malignant epithelial cells, similar to those previously reported in the mouse reproductive tract. Serial sections of malignant specimens show a good correlation between the localization of lactoferrin mRNA and protein in individual epithelial cells by in situ RNA hybridization and immunohistochemistry. Although the degree of lactoferrin expression in the adenocarcinomas did not correlate

with the tumor stage, grade, or depth of invasion in these 12 patients, there was a striking inverse correlation between the presence of progesterone receptors and lactoferrin in all 8 lactoferrin-positive adenocarcinomas. In summary, lactoferrin is expressed in a region of normal endometrium known as the zona basalis which is not shed with menstruation and is frequently overexpressed by progesterone receptor-negative cells in endometrial adenocarcinomas.

PMID: 7867003 [PubMed - indexed for MEDLINE]

1: Tumour Biol. 2004 Jul-Aug;25(4):161-71.

KARGER Full Text

Alteration of frizzled expression in renal cell carcinoma.

Janssens N, Andries L, Janicot M, Perera T, Bakker A.

Department of Biochemistry, University of Antwerp, Wilrijk, Belgium. njansse9@prdbe.jnj.com

To evaluate the involvement of frizzled receptors (Fzds) in oncogenesis, we investigated mRNA expression levels of several human Fzds in more than 30 different human tumor samples and their corresponding (matched) normal tissue samples, using real-time quantitative PCR. We observed that the mRNA level of Fzd5 was markedly increased in 8 of 11 renal carcinoma samples whilst Fzd8 mRNA was increased in 7 of 11 renal carcinoma samples. Western blot analysis of crude membrane fractions revealed that Fzd5 protein expression in the matched tumor/normal kidney samples correlated with the observed mRNA level. Wnt/beta-catenin signaling pathway activation was confirmed by the increased expression of a set of target genes. Using a kidney tumor tissue array, Fzd5 protein expression was investigated in a broader panel of kidney tumor samples. Fzd5 membrane staining was detected in 30% of clear cell carcinomas, and there was a strong correlation with nuclear cyclin D1 staining in the samples. Our data suggested that altered expression of certain members of the Fzd family, and their downstream targets, could provide alternative mechanisms leading to activation of the Wnt signaling pathway in renal carcinogenesis. Fzd family members may have a role as a biomarker.

PMID: 15557753 [PubMed - indexed for MEDLINE]

Display Abstract Show 20 Sort by Send to

May 22 2006 06:31:57

1: Breast Cancer Res Treat. 1992;24(1):71-4.

Related Articles, Links

Expression of the pS2 gene in breast tissues assessed by pS2-mRNA analysis and pS2-protein radioimmunoassay.

Hahnel E, Robbins P, Harvey J, Sterrett G, Hahnel R.

Department of Pathology, University of Western Australia, Queen Elizabeth II Medical Centre, Nedlands.

The expression of the pS2 gene in breast tissues was assessed by measuring pS2-protein using a radioimmunoassay, and by determining pS2-mRNA using Northern blotting. There was a good correlation between the two measurements, indicating that expression of the pS2 gene in breast tissues may be assessed by either method. Since radioimmunoassay is technically easier and more efficient than Northern blotting, radioimmunoassay will be the method of choice in routine applications.

PMID: 1463873 [PubMed - indexed for MEDLINE]

May 22 2006 06:31:57

1: Int J Oncol. 2005 Nov;27(5):1257-63.

Expression of human telomerase reverse transcriptase gene and protein, and of estrogen and progesterone receptors, in breast tumors: preliminary data from neo-adjuvant chemotherapy.

Kammori M, Izumiyama N, Hashimoto M, Nakamura K, Okano T, Kurabayashi R, Naoki H, Honma N, Ogawa T, Kaminishi M, Takubo K.

Division of Breast and Endocrine Surgery, Department of Surgery, Graduate School of Medicine, The University of Tokyo, Japan. kanmori-dis@umin.ac.jp

Human telomerase reverse transcriptase (hTERT), the catalytic subunit of telomerase, is very closely associated with telomerase activity. Telomerase has been implicated in cellular immortalization and carcinogenesis. In situ detection of hTERT will aid in determining the localization of telomerase-positive cells. The aim of this study was to detect expression of hTERT mRNA, hTERT protein, estrogen receptor (ER) and progesterone receptor (PR) in paraffin-embedded breast tissue samples and to investigate the relationship between hTERT expression and various clinicopathological parameters in breast tumorigenesis. We used in situ hybridization (ISH) to examine hTERT gene expression, and immunohistochemistry (IHC) to examine expression of hTERT protein, ER and PR, in breast tissues including 64 adenocarcinomas, 2 phyllode tumors and their adjacent normal breast tissues. hTERT gene expression was detected by ISH in 56 (88%) carcinomas, but in neither of the 2 phyllode tumors. hTERT protein expression was detected by IHC in 52 (81%) carcinomas, but in neither of the 2 phyllode tumors. Moreover, ER and PR were expressed in 42 (66%) and 42 (66%) carcinomas, respectively, and in neither of the 2 phyllode tumors. In 4 cases of breast carcinoma that strongly expressed hTERT gene and protein before treatment, neo-adjuvant chemotherapy led to disappearance of gene and protein expression in all cases. There was

a strong correlation between detection of hTERT gene expression by ISH and of hTERT protein by ICH in tissue specimens from breast tumors. These results suggest that detection of hTERT protein by ICH can be used to distinguish breast cancers as a potential diagnostic and therapeutic marker.

PMID: 16211220 [PubMed - indexed for MEDLINE]

[CANCER RESEARCH 61, 2409–2412, March 15, 2001]

Advances in Brief

BMI-1 Gene Amplification and Overexpression in Hematological Malignancies Occur Mainly in Mantle Cell Lymphomas[1]

Sílvia Beà, Frederic Tort, Magda Pinyol, Xavier Puig, Luis Hernández, Silvia Hernández, Pedro L. Fernández, Maarten van Lohuizen, Dolors Colomer, and Elias Campo[2]

The Hematopathology Section, Laboratory of Anatomic Pathology, Hospital Clinic, Institut d'Investigacions Biomèdiques "August Pi i Sunyer" (IDIBAPS), University of Barcelona, 08036 Barcelona, Spain [S. B., F. T., M. P., X. P., L. H., S. H., P. L. F., D. C., E. C.], and Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, 1066 CX Amsterdam, Netherlands [M. v. L.]

Abstract

The *BMI-1* gene is a putative oncogene belonging to the Polycomb group family that cooperates with c-*myc* in the generation of mouse lymphomas and seems to participate in cell cycle regulation and senescence by acting as a transcriptional repressor of the *INK4a/ARF* locus. The *BMI-1* gene has been located on chromosome 10p13, a region involved in chromosomal translocations in infant leukemias, and amplified in occasional non-Hodgkin's lymphomas (NHLs) and solid tumors. To determine the possible alterations of this gene in human malignancies, we have examined 160 lymphoproliferative disorders, 13 myeloid leukemias, and 89 carcinomas by Southern blot analysis and detected *BMI-1* gene amplification (3- to 7-fold) in 4 of 36 (11%) mantle cell lymphomas (MCLs) with no alterations in the *INK4a/ARF* locus. BMI-1 and $p16^{INK4a}$ mRNA and protein expression were also studied by real-time quantitative reverse transcription-PCR and Western blot, respectively, in a subset of NHLs. BMI-1 expression was significantly higher in chronic lymphocytic leukemia and MCL than in follicular lymphoma and large B cell lymphoma. The four tumors with gene amplification showed significantly higher mRNA levels than other MCLs and NHLs with the *BMI-1* gene in germline configuration. Five additional MCLs also showed very high mRNA levels without gene amplification. A good correlation between BMI-1 mRNA levels and protein expression was observed in all types of lymphomas. No relationship was detected between BMI-1 and $p16^{INK4a}$ mRNA levels. These findings suggest that *BMI-1* gene alterations in human neoplasms are uncommon, but they may contribute to the pathogenesis in a subset of malignant lymphomas, particularly of mantle cell type.

Introduction

The *BMI-1*[3] gene is a putative oncogene of the Polycomb group originally identified by retroviral insertional mutagenesis in Eμ-c-myc transgenic mice infected with the Moloney murine leukemia virus (1, 2). These animals had a rapid development of pre-B cell lymphomas showing frequent proviral insertions near the *BMI-1* gene. This integration resulted in BMI-1 overexpression suggesting a cooperative effect between *C-MYC* and *BMI-1* genes in the development of these tumors (3, 4). Recent studies have indicated that the *BMI-1* gene may also participate in cell cycle control and senescence through the *INK4a/ARF* locus by acting as an upstream negative regulator of $p16^{INK4a}$ and $p14/p19^{ARF}$ gene expression (5). The human *BMI-1* gene has been mapped to chromosome 10p13 (6), a region involved in chromosomal translocations in infant leukemias (7) and rearrangements in malignant T cell lymphomas (8, 9). More recently, high-level DNA amplifications of this region have been found by comparative genomic hybridization in NHLs and solid tumors (10, 11). However, the possible implication of the *BMI-1* gene in these alterations and its role in the pathogenesis of human tumors is not known. The aim of this study was to analyze the possible *BMI-1* gene alterations and expression in a large series of human neoplasms and to determine the relationship with *INK4a/ARF* locus aberrations.

Materials and Methods

Case Selection. A series of 262 human tumors, including 173 hematological malignancies and 89 carcinomas (Table 1), matched normal tissues from all carcinomas, 11 samples of normal peripheral mononuclear cells, and 5 reactive lymph nodes and tonsils, were selected based on the availability of frozen samples for molecular analysis.

DNA Extraction and Southern Blot Analysis. Genomic DNA was obtained using Proteinase K/RNase treatment. 15 μg were digested with *Eco*RI and *Hin*dIII restriction enzymes (Life Technologies, Inc., Gaithersburg, MD), for Southern blot analysis and hybridized with a 1.5-kb *Pst*I fragment of the partial *BMI-1* cDNA (6).

RNA Extraction and Real-time Quantitative RT-PCR. Total RNA was obtained from 67 lymphoid neoplasms (10 CLLs, 27 MCLs, 8 FLs, and 22 LCLs) using guanidine/isothiocyanate extraction and cesium/chloride gradient centrifugation. One μg of total RNA was transcribed into cDNA using MMLV-reverse transcriptase (Life Technologies, Inc.) and random hexamers, following manufacturer's directions. Sequences of the *BMI-1* and the *p16* detection probes and primers were designed using the Primer Express program (Applied Biosystems, Foster City) as follows: *BMI-1* sense, 5′-CTGGTTGCCATTGACAGC-3′; *BMI-1* antisense, 5′-CAGAAAATGAATGCGAGCCA-3′; *p16* sense, 5′-CAACGCACCGAATAGTTACGG-3′; *p16* antisense, 5′-AACTTCGTCCTCCAGAGTCGC-3′. The probes *BMI-1*, 5′-CAGCTCGCTTCAAGATGGCCGC-3′, and *p16*, 5′-CGGAGGCCGATCCAGGTGGGTA-3′, were labeled with 6-carboxy-fluorescein as the reporter dye. The TaqMan-GAPDH Control Reagents (Applied Biosystems) were used to amplify and detect the *GAPDH* gene, as recommended by the manufacturer. The quantitative assay amplified 1 μml of cDNA in two to four replicates using the primers and probes described above and the standard master mix (Applied Biosystems). All reactions were performed in an ABI PRISM 7700 Sequence Detector System (Applied Biosystems). GAPDH, BMI-1, and $p16^{INK4a}$ expression was related to a standard curve derived from serial dilutions of Raji cDNA. The RUs of BMI-1 and $p16^{INK4a}$ expression were defined as the mRNA levels of these genes normalized to the GADPH expression level in each case.

Protein Analysis. Whole-cell protein extracts were obtained from additional frozen tissue available in 31 cases (7 CLLs, 12 MCLs, 8 FLs, and 4 LCLs), loaded onto a 10% SDS-polyacrylamide gel, and electroblotted to a nitrocellulose membrane (Amersham). Blocked membranes were incubated sequentially with the monoclonal antibody BMI-F6 (12), antimouse conju-

Received 10/16/00; accepted 1/29/01.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported by Grant SAF 99/20 from Comision Interministerial de Ciencia y Tecnologia, European Union Contract QLG1-CT-2000-689, the Asociación Española contra el Cáncer, and Generalitat de Catalunya 98SGR21. S. B. and F. T. were fellows supported by Spanish Ministerio de Educación y Cultura, and S. H. was supported by the Asociación Española contra el Cáncer and the Fundació Rius i Virgili.

[2] To whom requests for reprints should be addressed, at the Department of Pathology, Hospital Clinic, University of Barcelona, Villarroel 170, 08036-Barcelona, Spain. Phone: 34 93 227 5450; Fax: 34 93 227 5572; E-mail: campo@medicina.ub.es.

[3] The abbreviations used are: BMI-1, B cell-specific Moloney murine leukemia virus integration site 1; NHL, non-Hodgkin's lymphoma; CLL, chronic lymphocytic leukemia; FL, follicular lymphoma; LCL, large B cell lymphoma; MCL, mantle cell lymphoma; RT-PCR, reverse-transcription-PCR; RU, relative units.

Table 1 *Hematological malignancies and solid tumor samples analyzed for BMI-1 gene alterations*

Tissue samples	No. of cases
Hematological malignancies	
Hodgkin's disease	2
B cell lymphoproliferative disorders	
B-Acute lymphoblastic leukemia	14
CLL	29
Hairy cell leukemia	4
FL	15
MCL	36
LCL	40
T cell lymphoproliferative disorders	
T-Acute lymphoblastic leukemia	8
Large granular cell leukemia	4
Peripheral T-cell lymphoma	8
Myeloproliferative disorders	
Acute myeloid leukemia	7
Chronic myeloid leukemia	6
Solid tumors	
Colon carcinoma	26
Breast carcinoma	29
Laryngeal squamous cell carcinoma	34
Total	262

gated to horseradish peroxidase (Amersham), and detected by enhanced chemiluminescence (Amersham) according to the manufacturer's recommendations.

Statistical Analysis. Because of the non-normal distribution of the samples and the small size of some subsets of tumors, the statistical evaluation was performed using nonparametric tests (SPSS, version 9.0). Comparison between mRNA expression levels in the different groups of NHLs was performed using the Kruskal-Wallis Test, with a P for significance set at 0.05. For differences between particular groups, the conservative Bonferroni procedure was performed, and the P was set at 0.005. The remaining statistical analyses were carried out using the Mann-Whitney nonparametric U test (significance, $P < 0.05$). The comparison between BMI-1 and $p16^{INK4a}$ quantitative mRNA levels was also performed using the Pearson's correlation coefficient.

Results

***BMI-1* Gene Amplification.** The *BMI-1* gene was examined by Southern blot in a large series of human tumors and normal samples (Table 1). The cDNA probe used in the study detected three *Eco*RI fragments of 7.3, 3.8, and 2.6 kb and three *Hin*dIII fragments of 6.2, 4, and 3.5 kb. *BMI-1* gene amplification (3- to 7-fold) was detected in 4 of 36 (11%) MCLs (Fig. 1). The amplifications were confirmed with both restriction enzymes. The amplified MCLs were two blastoid and two typical variants. No amplifications were observed in any of the solid tumors when compared with their respective matched non-neoplastic mucosa. No *BMI-1* gene rearrangements were observed in any of the samples examined.

BMI-1 mRNA Expression. To determine the BMI-1 expression pattern in NHL we analyzed BMI-1 mRNA levels by real-time quantitative RT-PCR in 67 lymphomas (10 CLLs, 27 MCLs, 8 FLs, and 22 LCLs), including the four tumors with gene amplification. A distinct BMI-1 mRNA expression pattern was observed in the different types of lymphomas (Fig. 2; Kruskal-Wallis Test; $P < 0.001$). The BMI mRNA levels in CLLs (mean, 2.2 RU; SD, 1.3) and MCLs with no *BMI-1* gene amplification (mean, 2.5 RU; SD, 2.3) were significantly higher than in FLs (mean, 0.9 RU; SD, 0.8) and LCLs (mean, 0.6 RU; SD, 0.4; Mann-Whitney nonparametric U test; $P < 0.01$). The 4 MCLs with *BMI-1* gene amplification showed significantly higher levels of expression than all other groups of tumors (mean, 5.1 RU; SD, 1.6; $P < 0.005$). In addition, five typical MCLs with no structural alterations of the gene also showed very high levels of BMI-1 mRNA expression ranging from 4 to 9.8 RU, similar to cases with gene amplification (Fig. 2*A*).

BMI-1 Protein Expression. BMI-1 protein expression was examined by Western blot in 31 tumors (7 CLLs; 12 MCLs, including two cases with *BMI-1* gene amplification and 4 cases with mRNA overexpression and no structural alteration of the gene; 8 FLs, and 4 LCLs) in which additional frozen tissue was available. The monoclonal antibody against BMI-1 detected three closely migrating proteins of M_r 45,000–48,000 (2). The two more slowly migrating bands probably represent phosphorylated isoforms of the protein (12). The two MCLs with gene amplification and three of four cases with mRNA overexpression without amplification of the gene showed very high levels of protein expression. The remaining MCLs and CLLs showed intermediate levels of protein expression, whereas low- or no-expression signals were detected in the LCLs and FLs included in the study (Fig. 3). These results indicate that BMI-1 protein expression in NHL is concordant with the mRNA levels observed by real-time quantitative RT-PCR.

Relationship between BMI-1 and $p16^{INK4a}$ Gene Alterations. The *INK4a/ARF* locus has been recently identified as a downstream target of the transcriptional repressing activity of the *BMI-1* gene, suggesting that this gene may contribute to human neoplasias with wild type *INK4/ARF* (5). Most of the lymphoproliferative disorders analyzed in the present study, including the four cases with *BMI-1* gene amplification, had been previously examined for *p53* gene mutations and *INK4a/ARF* locus alterations, including gene deletions, mutations, hypermethylation, and expression (13, 14). The four MCLs with *BMI-1* gene amplification and mRNA overexpression and the five tumors with BMI-1 mRNA overexpression with no structural alterations of the gene showed a wild-type configuration of the *INK4a/ARF* locus (13). However, one case with *BMI-1* gene amplification and one case with mRNA overexpression with no alteration of the gene showed *p53* gene mutations associated with allelic deletions.

To determine the possible relationship between BMI-1 and $p16^{INK4a}$ mRNA expression, $p16^{INK4a}$ mRNA levels were evaluated by real-time quantitative RT-PCR in 50 tumors (10 CLLs, 27 MCLs, and 13 LCLs), including 6 cases with alterations in the *INK4a/ARF* locus (2 MCLs and 1 LCL with $p16^{INK4a}$ gene deletion, 2 LCLs with *p16* promoter hypermethylation, and 1 CLL with $p16^{INK4a}$ gene mutation), and the 4 lymphomas with *BMI-1* amplification. Negative or negligible levels of $p16^{INK4a}$ were observed in the 6 tumors with *INK4a/ARF* locus alterations. These cases were not included in the comparisons between BMI-1 and $p16^{INK4a}$ mRNA expression. The $p16^{INK4a}$ expression levels were relatively similar in the different types of tumors. Only LCLs tended to have lower levels of expression, but the differences did not reach statistical significance (Fig. 2*B*). No differences were observed in the $p16^{INK4a}$ mRNA levels between tumors with *BMI-1* gene amplification and overexpression and lymphomas with germline configuration of the gene.

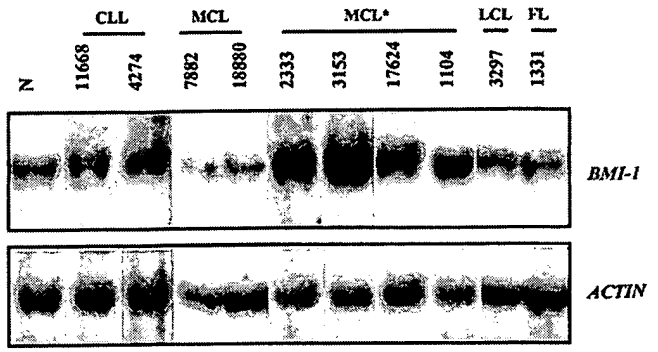


Fig. 1. Southern blot analysis of *BMI-1* gene. Four MCLs (*MCL**) showed *BMI-1* gene amplification (3- to 7-fold) compared with non-neoplastic tissues (*N*) and other NHLs. No amplifications or gene rearrangements were detected in the remaining NHLs and carcinomas included in the study.

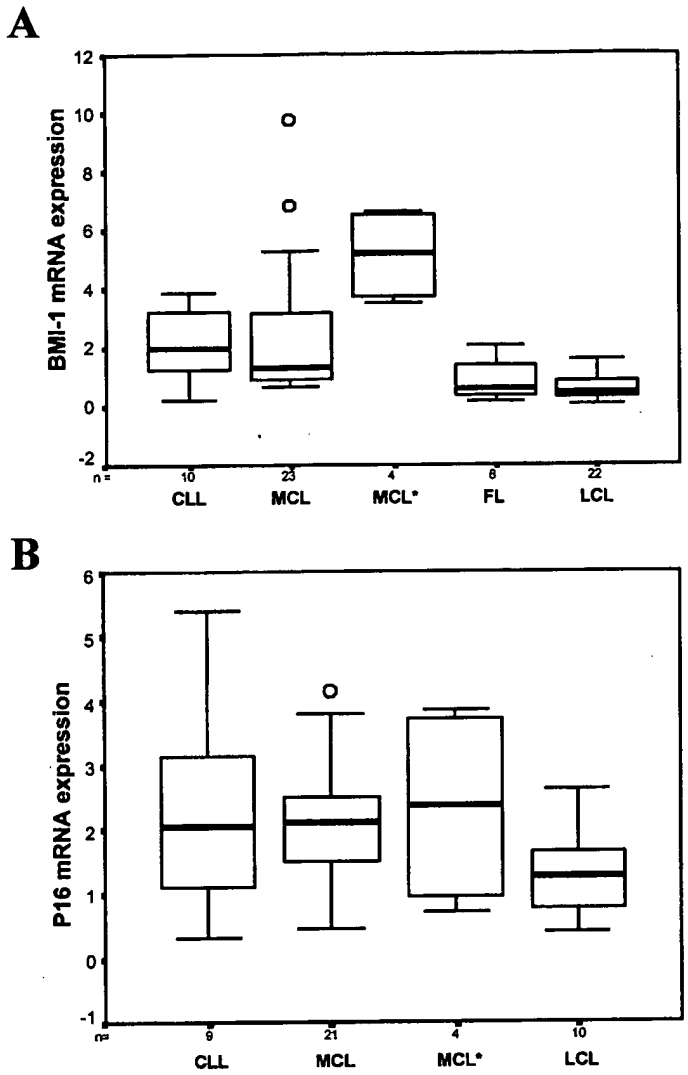


Fig. 2. *A*, quantitative BMI-1 mRNA transcript analysis (median and range) using real-time RT-PCR in a series of NHLs. MCLs with *BMI-1* gene amplification (*MCL**) revealed significantly higher overall BMI-1 mRNA levels than all other types of NHLs, including MCLs with no structural alterations of the gene ($P < 0.005$). MCLs and CLLs expressed significantly higher levels than FLs and LCLs ($P < 0.001$). Results are depicted as the ratio of absolute BMI-1:GADPH mRNA transcript numbers (RU). *Bars*, SD. *B*, quantitative $p16^{INK4a}$ mRNA transcript analysis (median and range) using real-time RT-PCR in a series of NHLs. Expression levels were relatively similar in the different types of tumors. Results are depicted as the ratio of absolute $p16^{INK4a}$:GADPH mRNA transcript numbers (RU). *Bars*, SD.

Discussion

In the present study, we have examined a large series of human tumors for the presence of gene alterations and mRNA expression of the *BMI-1* gene. Gene amplification was identified in four MCLs. These tumors showed significantly higher levels of mRNA and protein expression compared with other lymphomas with *BMI-1* in germline configuration. BMI-1 expression levels were also highly upregulated in a subset of MCLs with no apparent structural alterations of the gene. No alterations were detected in any of the different types of carcinomas included in the study. *BMI-1* is considered an oncogene belonging to the Polycomb group family of genes. These proteins mainly act as transcriptional regulators, controlling specific target genes involved in development, cell differentiation, proliferation, and senescence. Different studies have shown the implication of BMI-1 overexpression in the development of lymphomas in murine and feline animal models (3, 4). The findings of the present study indicate for the first time that *BMI-1* gene alterations in human neoplasms are an uncommon phenomenon, but they seem to occur mainly in a subset of NHLs, particularly of mantle cell type.

The human *BMI-1* gene has been mapped to chromosome 10p13. High-level DNA amplifications and gains in this region have been identified by comparative genomic hybridization in occasional solid tumors and NHLs (10, 11). Different chromosomal translocations involving the 10p13 region have also been identified in infant leukemias and T cell lymphoproliferative disorders (7, 8, 15). Most acute leukemias with this chromosomal alteration occur in children <12 months of age, whereas it seems to be extremely rare in adults. 10p translocations in T-cell lymphoproliferative disorders have been observed mainly in adult T cell leukemia/lymphomas and occasional cutaneous T cell lymphomas. In our study, we did not observe *BMI-1* rearrangements or amplifications in any of the acute leukemias or T cell lymphomas. However, all of the acute leukemias in this study were diagnosed in patients over 16 years, and no adult T cell leukemia/lymphomas or cutaneous lymphomas could be included in the series. Similarly, high-level DNA amplifications at the 10p13 region have been detected in head and neck carcinomas and other solid tumors. Although we found no evidence for *BMI-1* gene rearrangements or amplifications in a substantial set of carcinomas, this does not exclude the possibility of increased gene expression or protein levels in these tumors. Additional studies are required to elucidate the possible involvement of *BMI-1* in these particular groups of human neoplasms.

In human hematopoietic cells, BMI-1 is preferentially expressed in primitive CD34+ bone marrow cells, whereas it is negative or very low in more mature CD34− cells (16). In peripheral lymphocytes, and particularly in follicular B cells, BMI-1 protein expression has been detected in resting cells of the mantle zone, whereas it is downregulated in proliferating germinal center cells (17, 18). These observations indicate that BMI-1 expression in normal hematopoietic cells is tightly regulated in relation with cell differentiation in bone marrow and antigen-specific response in peripheral lymphocytes. BMI-1 expression in human tumors has not been examined previously. In this study, we have demonstrated that BMI-1 mRNA and protein expression show a distinct pattern in different types of lymphomas. Thus, BMI-1 levels were low in LCLs and FLs and significantly higher in MCLs and CLLs. These findings suggest that BMI-1 expression patterns in B cell lymphomas maintain in part the expression profile of their normal cell counterparts; because FLs and at least a subgroup of LCLs are considered lymphomas derived from follicular germinal center cells, whereas MCLs and CLLs are tumors mainly derived from naive pregerminal center cells. However, the four MCLs with *BMI-1* gene amplification expressed significantly higher mRNA levels than all other tumors. In addition, five MCLs with no structural alterations of the gene showed high mRNA levels similar to those observed in tumors with *BMI-1* gene amplification, suggesting that other mechanisms may be involved in up-regulation of the gene in these lymphomas. Different studies using animal models have shown a dose-dependent effect of *BMI-1* gene expression on skeleton development

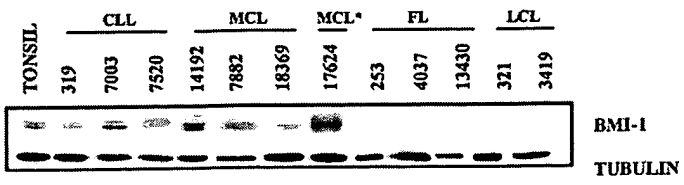


Fig. 3. Western blot analysis of BMI-1 protein in NHLs. The amplified MCL (17624) showed the highest BMI-1 protein levels, whereas other MCLs and CLLs had intermediate levels of expression. Very low or negative signal was observed in FLs and LCLs.

and lymphomagenesis (1, 3). These observations suggest that the high mRNA and protein levels detected in a subset of MCLs may play a role in the pathogenesis of these neoplasms.

Recent studies have identified the *INK4/ARF* locus as a downstream target of the BMI-1 transcriptional repressor activity, suggesting that BMI-1 overexpression may contribute to human neoplasias that retain the wild-type *INK4a/ARF* locus (5). Interestingly, in our study, *BMI-1* amplification and overexpression appeared in tumors with no alterations in $p16^{INK4a}$ and $p14^{ARF}$ genes. However, we could not detect differences in the expression levels of $p16^{INK4a}$ in tumors with and without *BMI-1* gene alterations. The reasons for this apparent discrepancy with experimental observations are not clear. One possibility may be that genes other than *INK4a/ARF* are the main targets of BMI-1 repressor activity in these tumors. Particularly, different genes of the HOX family are regulated by BMI-1 and may also be involved in lymphomagenesis (19, 20).

In conclusion, the findings of this study indicate that *BMI-1* gene expression is differentially regulated in B cell lymphomas. Alterations of the gene seem to be an uncommon phenomenon in human neoplasms, but they may contribute to the pathogenesis in a subset of MCLs. Although, *BMI-1* gene alterations occurred in tumors with wild-type *INK4a/ARF* locus, the possible cooperation between these genes and the oncogenic mechanisms of BMI-1 in human neoplasms require additional analysis.

Acknowledgments

The authors thank Iracema Nayach for her excellent technical assistance.

References

1. Haupt, Y., Alexander, W. S., Barri, G., Klinken, S. P., and Adams, J. M. Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in E μ-myc transgenic mice. Cell, *65:* 753–763, 1991.
2. van Lohuizen, M., Verbeek, S., Scheijen, B., Wientjens, E., van der Gulden, H., and Berns, A. Identification of cooperating oncogenes in E μ-myc transgenic mice by provirus tagging. Cell, *65:* 737–752, 1991.
3. Alkema, M. J., Jacobs, H., van Lohuizen, M., and Berns, A. Perturbation of B and T cell development and predisposition to lymphomagenesis in EμBmi1 transgenic mice require the Bmi1 RING finger. Oncogene, *15:* 899–910, 1997.
4. Haupt, Y., Bath, M. L., Harris, A. W., and Adams, J. M. *Bmi-1* transgene induces lymphomas and collaborates with myc in tumorigenesis. Oncogene, *8:* 3161–3164, 1993.
5. Jacobs, J. J., Kieboom, K., Marino, S., DePinho, R. A., and van Lohuizen, M. The oncogene and Polycomb-group gene *bmi-1* regulates cell proliferation and senescence through the *ink4a* locus. Nature (Lond.), *397:* 164–168, 1999.
6. Alkema, M. J., Wiegant, J., Raap, A. K., Berns, A., and van Lohuizen, M. Characterization and chromosomal localization of the human proto-oncogene *BMI-1*. Hum. Mol. Genet., *2:* 1597–1603, 1993.
7. Pui, C. H., Raimondi, S. C., Murphy, S. B., Ribeiro, R. C., Kalwinsky, D. K., Dahl, G. V., Crist, W. M., and Williams, D. L. An analysis of leukemic cell chromosomal features in infants. Blood, *69:* 1289–1293, 1987.
8. Berger, R., Baranger, L., Bernheim, A., Valensi, F., Flandrin, G., and Berheimm, A. T. Cytogenetics of T-cell malignant lymphoma. Report of 17 cases and review of the chromosomal breakpoints. Cancer Genet. Cytogenet., *36:* 123–130, 1988.
9. D'Alessandro, E., Paterlini, P., Lo Re, M. L., Di Cola, M., Ligas, C., Quaglino, D., and Del Porto, G. Cytogenetic follow-up in a case of Sezary syndrome. Cancer Genet. Cytogenet., *45:* 231–236, 1990.
10. Bea, S., Ribas, M., Hernandez, J. M., Bosch, F., Pinyol, M., Hernandez, L., Garcia, J. L., Flores, T., Gonzalez, M., Lopez-Guillermo, A., Piris, M. A., Cardesa, A., Montserrat, E., Miro, R., and Campo, E. Increased number of chromosomal imbalances and high-level DNA amplifications in mantle cell lymphoma are associated with blastoid variants. Blood, *93:* 4365–4374, 1999.
11. Knuutila, S., Bjorkqvist, A. M., Autio, K., Tarkkanen, M., Wolf, M., Monni, O., Szymanska, J., Larramendy, M. L., Tapper, J., Pere, H., el-Rifai, W., Hemmer, S., Wasenius, V. M., Vidgren, V., and Zhu, Y. DNA copy number amplifications in human neoplasms: review of comparative genomic hybridization studies. Am. J. Pathol., *152:* 1107–1123, 1998.
12. Alkema, M. J., Bronk, M., Verhoeven, E., Otte, A., van't Veer, L. J., Berns, A., and van Lohuizen, M. Identification of Bmi1-interacting proteins as constituents of a multimeric mammalian polycomb complex. Genes Dev., *11:* 226–240, 1997.
13. Pinyol, M., Hernández, L., Martínez, A., Cobo, F., Hernández, S., Beà, S., López-Guillermo, A., Nayach, I., Palacín, A., Nadal, A., Fernández, P., Montserrat, E., Cardesa, A., and Campo, E. *INK4a/ARF* locus alterations in human non-Hodgkin's lymphomas mainly occur in tumors with wild type *p53* gene. Am. J. Pathol., *156:* 1987–1996, 2000.
14. Pinyol, M., Cobo, F., Beà, S., Jares, P., Nayach, I., Fernández, P. L., Montserrat, E., Cardesa, A., and Campo, E. *p16INK4a* gene inactivation by deletions, mutations, and hypermethylation is associated with transformed and aggressive variants of non-Hodgkin's lymphomas. Blood, *91:* 2977–2984, 1998.
15. Foot, A. B., Oakhill, A., and Kitchen, C. Acute monoblastic leukemia of infancy in Klinefelter's syndrome. Cancer Genet. Cytogenet., *61:* 99–100, 1992.
16. Lessard, J., Baban, S., and Sauvageau, G. Stage-specific expression of polycomb group genes in human bone marrow cells. Blood, *91:* 1216–1224, 1998.
17. Raaphorst, F. M., van Kemenade, F. J., Fieret, E., Hamer, K. M., Satijn, D. P., Otte, A. P., and Meijer, C. J. Cutting edge: polycomb gene expression patterns reflect distinct B cell differentiation stages in human germinal centers. J. Immunol., *164:* 1–4, 2000.
18. Raaphorst, F. M., van Kemenade, F. J., Blokzijl, T., Fieret, E., Hamer, K. M., Satijn, D. P., Otte, A. P., and Meijer, C. J. Coexpression of *BMI-1* and *EZH2* polycomb group genes in Reed-Sternberg cells of Hodgkin's disease. Am. J.. Pathol., *157:* 709–715, 2000.
19. Gould, A. Functions of mammalian Polycomb group and trithorax group related genes. Curr. Opin. Genet. Dev., *7:* 488–494, 1997.
20. van Oostveen, J., Bijl, J., Raaphorst, F., Walboomers, J., and Meijer, C. The role of homeobox genes in normal hematopoiesis and hematological malignancies. Leukemia, *13:* 1675–1690, 1999.

American Journal of Pathology, Vol. 155, No. 3, September 1999

Id-1 and Id-2 Are Overexpressed in Pancreatic Cancer and in Dysplastic Lesions in Chronic Pancreatitis

Haruhisa Maruyama,* Jörg Kleeff,* Stefan Wildi,* Helmut Friess,† Markus W. Büchler,† Mark A. Israel,‡ and Murray Korc*

From the Division of Endocrinology, Diabetes, and Metabolism, Departments of Medicine, Biological Chemistry and Pharmacology, University of California, Irvine, California; the Department of Visceral and Transplantation Surgery,† University of Bern, Bern, Switzerland; and the Preuss Laboratory,‡ Department of Neurological Surgery, University of California, San Francisco, California*

Id proteins antagonize basic helix-loop-helix proteins, inhibit differentiation, and enhance cell proliferation. In this study we compared the expression of Id-1, Id-2, and Id-3 in the normal pancreas, in pancreatic cancer, and in chronic pancreatitis (CP). Northern blot analysis demonstrated that all three Id mRNA species were expressed at high levels in pancreatic cancer samples by comparison with normal or CP samples. Pancreatic cancer cell lines frequently coexpressed all three Ids, exhibiting a good correlation between Id mRNA and protein levels, as determined by immunoblotting with highly specific anti-Id antibodies. Immunohistochemistry using these antibodies demonstrated the presence of faint Id-1 and Id-2 immunostaining in pancreatic ductal cells in the normal pancreas, whereas Id-3 immunoreactivity ranged from weak to strong. In the cancer tissues, many of the cancer cells exhibited abundant Id-1, Id-2, and Id-3 immunoreactivity. Scoring on the basis of percentage of positive cells and intensity of immunostaining indicated that Id-1 and Id-2 were increased significantly in the cancer cells by comparison with the respective controls. Mild to moderate Id immunoreactivity was also seen in the ductal cells in the CP-like areas adjacent to these cells and in the ductal cells of small and interlobular ducts in CP. In contrast, in dysplastic and atypical papillary ducts in CP, Id-1 and Id-2 immunoreactivity was as significantly elevated as in the cancer cells. These findings suggest that increased Id expression may be associated with enhanced proliferative potential of pancreatic cancer cells and of proliferating or dysplastic ductal cells in CP. *(Am J Pathol 1999, 155:815–822)*

Basic helix-loop-helix (bHLH) proteins play an important role as transcription factors in cellular development, proliferation, and differentiation.[1,2] The basic domain of the bHLHs is required for binding to an E-box DNA sequence, thus promoting transcription of specific target genes. The HLH domain promotes dimer formation with various members of the bHLH protein family.[1,2] Homodimers of the class B family of bHLH proteins, including *Myo*D, *Neuro*D, and numerous other proteins, are known to activate tissue-specific genes.[3–5] These tissue-specific bHLHs typically form heterodimers with widely expressed class A bHLHs, which include proteins encoded by E2A, E2–2, HEB, and other genes (also termed E-proteins).[6–9] These heterodimers activate transcription of genes that are associated with differentiation.

Id genes encode a family of four HLH proteins that lack the basic DNA binding domain.[1,10] They act as dominant-negative HLH proteins by forming high affinity heterodimers with other bHLH proteins, thereby preventing them from binding to DNA and inhibiting transcription of differentiation-associated genes.[10–12] Id gene expression is down-regulated on differentiation in many cell types *in vitro* and *in vivo*.[13–18] In addition, Id proteins seem to be required for cell cycle progression through G_1/S phase in certain cell types, and interaction between Id-2 and pRB is associated with enhanced proliferation in some cell lines *in vitro*.[19–23]

Pancreatic cancer is the fifth leading cause of cancer death in the United States, with a mortality rate that virtually equals its incidence rate.[24] This malignancy is often associated with the overexpression of a variety of mitogenic growth factors and their receptors, and by oncogenic mutations of K-*ras* and inactivation of the p53 tumor suppressor gene.[25] We have recently reported that pancreatic cancers overexpress the HLH protein Id-2, and that enhanced expression of this protein is evident in the cytoplasm of the cancer cells within the pancreatic tumor mass.[26] It is not known, however, whether the expression of other Id proteins is altered in this malignancy, or whether their expression is altered in chronic pancreatitis

Contract grant sponsor: National Cancer Institute. Contract grant number: U. S. Public Health Service grant CA-40162.

Accepted for publication May 24, 1999.

Address reprint requests to Dr. Murray Korc, Division of Endocrinology, Diabetes and Metabolism, Medical Sciences I, C240, University of California, Irvine, CA 92697. E-mail: mkorc@uci.edu.

(CP), an inflammatory disease that is characterized by dysplastic ducts, foci of proliferating ductal cells, acinar cell degeneration, and fibrosis.[27] We now report that there is a five- to sixfold increase in Id-1 and Id-2 mRNA levels and a twofold increase in Id-3 mRNA levels in pancreatic cancer by comparison with the normal pancreas. In contrast, overall Id mRNA levels are not increased in CP.

Patients and Methods

Normal human pancreatic tissue samples from 7 male and 5 female donors (median age 41.8 years, range 14–68 years), CP tissues from 13 males and 1 female (median age 42.1 years; range 30–56 years), and pancreatic cancer tissues from 10 male and 6 female donors (median age 62.6 years; range 53–83 years) were obtained through an organ donor program and from surgical specimens from patients with severe symptomatic chronic pancreatitis or pancreatic cancer. A partial duodenopancreatectomy (Whipple/pylorus-preserving Whipple; $n = 13$), a left resection of the pancreas ($n = 2$), or a total pancreatectomy ($n = 1$) were carried out in the pancreatic cancer patients. According to the TNM classification of the Union Internationale Contre le Cancer (UICC) 6 tumors were stage 1, 1 was stage 2, and 9 were stage 3 ductal cell adenocarcinoma. Freshly removed tissue samples were fixed in 10% formaldehyde solution for 12 to 24 hours and paraffin-embedded for histological analysis. In addition, tissue samples were frozen in liquid nitrogen immediately on surgical removal and maintained in −80°C until use for RNA extraction. All studies were approved by the Ethics Committee of the University of Bern, Bern, Switzerland, and by the Human Subjects Committee at the University of California, Irvine, California.

Northern Blot Analysis

Northern blot analysis was carried out as described previously.[26,28] Briefly, total RNA was extracted by the single step acid guanidinium thiocyanate phenol chloroform method. RNA was size-fractionated on 1.2% agarose/1.8 mol/L formaldehyde gels, electrotransferred onto nylon membranes, and cross-linked by UV irradiation. Blots were prehybridized and hybridized with cDNA probes and washed under high stringency conditions. The following cDNA probes were used: a 979-bp human Id-1 cDNA probe, a 440-bp human Id-2 cDNA probe, and a 450-bp human Id-3 cDNA probe, covering the entire coding regions of Id-1, Id-2, and Id-3, respectively. A *Bam*HI 190-bp fragment of mouse 7S cDNA that hybridizes with human cytoplasmic RNA was used to confirm equal RNA loading and transfer. Blots were then exposed at −80°C to Kodak BioMax-MS films and the resulting autoradiographs were scanned to quantify the intensity of the radiographic bands.[26,28] For each sample the ratio of Id mRNA expression to 7S expression was calculated. To compare the relative increase in expression of the respective Id mRNA species in the cancer and CP samples, the same normal samples were used for normal/cancer and normal/CP membranes. The median score for Id-1, Id-2, and Id-3 mRNA levels in these normal samples was set to 100. Statistical analysis was performed with SigmaStat software (Jandel Scientific, San Raphael, CA). The rank sum test was used, and $P < 0.05$ was taken as the level of significance.

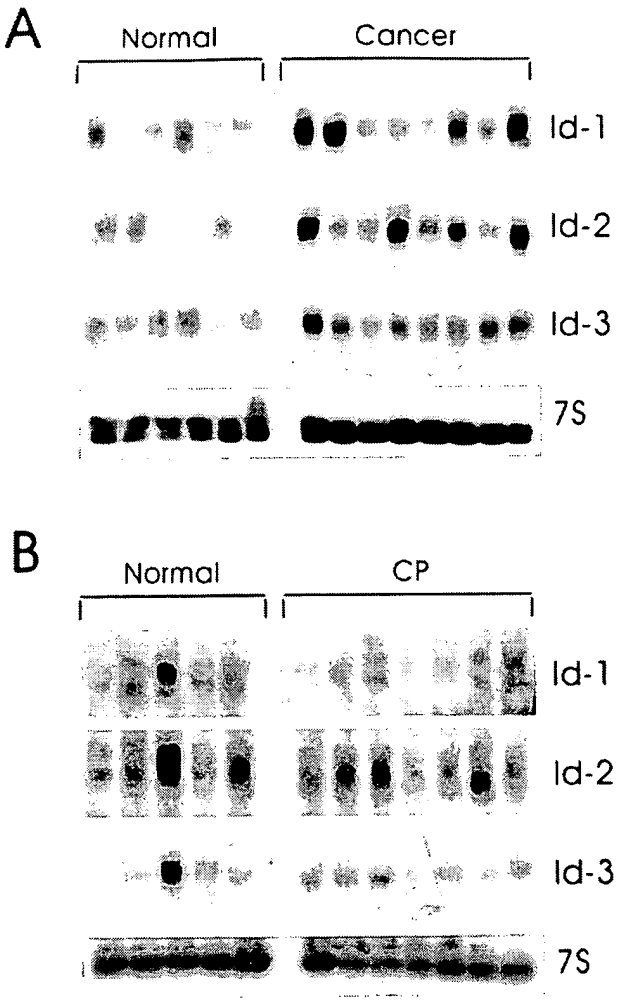


Figure 1. mRNA expression of Id-1, Id-2, and Id-3 in pancreatic cancer and chronic pancreatitis. Total RNA (20 μg/lane) from six normal, eight cancerous, and seven chronic pancreatitis tissue samples were subjected to Northern blot analysis using ^{32}P-labeled cDNA probes (500,000 cpm/ml) specific for Id-1, Id-2, and Id-3, respectively. A 7S cDNA probe (50,000 cpm/ml) was used as a loading and transfer control. Exposure times of the normal/cancer blots were 1 day for all Id probes, and 2 days for the normal/CP blots. Exposure time was 4 hours for mouse 7S cDNA. By comparison with the normal samples, Id-1 and Id-3 mRNA levels were elevated in 8 and 9 cancer samples, respectively, whereas Id-2 was elevated in 6 cancer samples.

Cell Culture and Western Blot Analysis

PANC-1, MIA-PaCa-2, ASPC-1, and CAPAN-1 human pancreatic cell lines were obtained from ATCC (Manassas, VA). COLO-357 human pancreatic cells were a gift from Dr. R. S. Metzger (Durham, NC). Cells were routinely grown in DMEM (COLO-357, MIA-PaCa-2, PANC-1) or RPMI (ASPC-1, CAPAN-1) supplemented with 10% fetal bovine serum, 100 U/ml penicillin, and 100 μg/ml streptomycin. For immunoblot analysis, exponentially growing

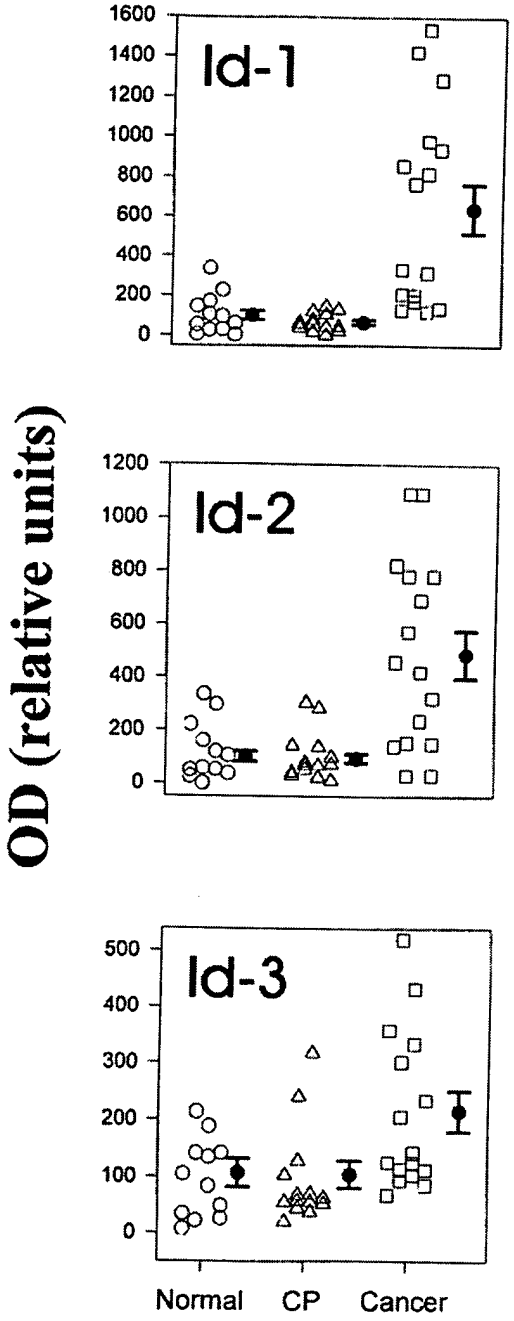


Figure 2. Densitometric analysis of Northern blots. Autoradiographs of Northern blots from 12 normal, 14 CP, and 16 pancreatic cancers were analyzed by densitometry. mRNA levels were determined by calculating the ratio of the optical density for the respective Id mRNA species in relation to the optical density of mouse 7S cDNA. To compare the relative increase in expression of the respective Id mRNA species in the cancer and CP samples, the same normal samples were used for normal/cancer and normal/CP membranes. Normal pancreatic tissues are indicated by circles, CP tissues by triangles, and cancer tissues by squares. Data are expressed as median scores ± SD. By comparison with the normal samples, only the cancer samples exhibited significant increases: 6.5-fold ($P < 0.01$) for Id-1, fivefold ($P < 0.01$) for Id-2, and twofold ($P = 0.027$) for Id-3.

Figure 3. Id mRNA and protein expression in pancreatic cancer cell lines. **Upper panels:** Total RNA (20 μg/lane) from 5 pancreatic cancer cell lines were subjected to Northern blot analysis using ^{32}P-labeled cDNA probes (500,000 cpm/ml) specific for Id-1, Id-2, and Id-3, respectively. Exposure times were 1 day for all Id probes. **Lower panels:** Immunoblotting. Cell lysates (30 μg/lane) were subjected to SDS-PAGE. Membranes were probed with specific Id-1, Id-2, and Id-3 antibodies. Visualization was performed by enhanced chemiluminescence.

cells (60–70% confluent) were solubilized in lysis buffer containing 50 mmol/L Tris-HCl, pH 7.4, 150 mmol/L NaCl, 1 mmol/L EDTA, 1 μg/ml pepstatin A, 1 mmol/L phenylmethylsulfonyl fluoride (PMSF), and 1% Triton X-100. Proteins were subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), transferred to Immobilon P membranes, and incubated for 90 minutes with the indicated antibodies and for 60 minutes with secondary antibodies against rabbit IgG. Visualization was performed by enhanced chemiluminescence.

Immunohistochemistry

Specific rabbit anti-human Id-1 (C-20), Id-2 (C-20), and Id-3 (C-20; all from Santa Cruz Biotechnology, Santa

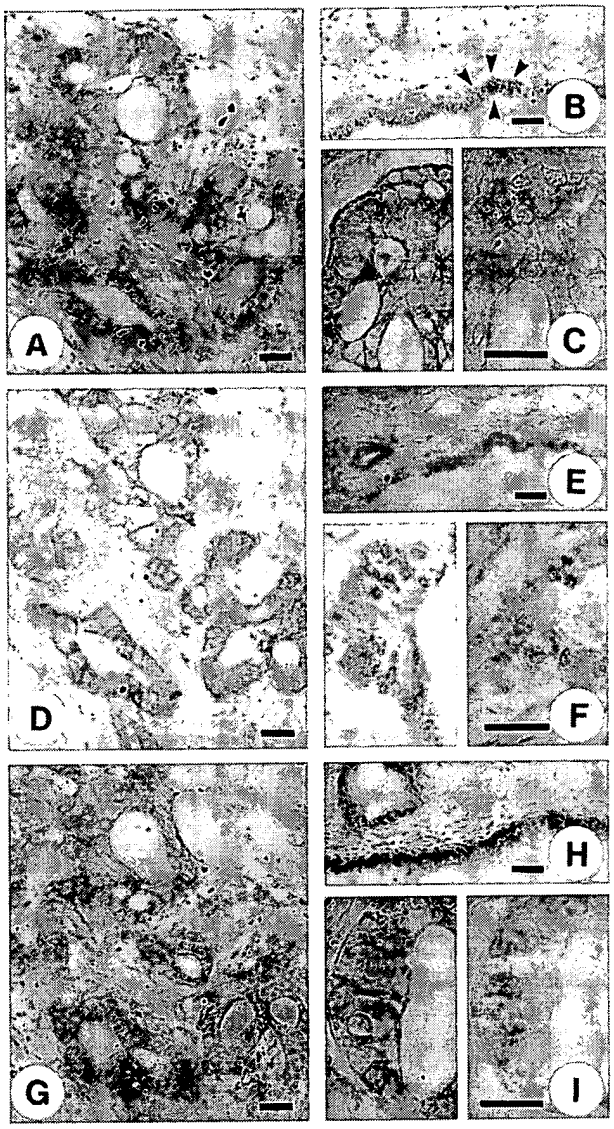


Figure 4. Normal and cancerous pancreatic tissues were subjected to immunostaining using highly specific anti-Id-1 (A-C), anti-Id-2 (D-F), and anti-Id-3 (G-I) antibodies as described in the Methods section. Moderate to strong Id-1 immunoreactivity was present in the cytoplasm of duct-like cancer cells (A and C, left panel). In the normal pancreas there was weak Id-1 immunoreactivity in the ductal cells (B). Preabsorption with the Id-1-specific blocking peptide abolished the Id-1 immunoreactivity (C, right panel). Strong Id-2 immunoreactivity was observed in the cytoplasm of the cancer cells that exhibited duct-like structures (D and F, left panel), whereas in the normal pancreas, there was only weak Id-2 immunoreactivity in the ductal cells (E). Preabsorption with the Id-2-specific blocking peptide abolished the Id-2 immunoreactivity (F, right panel). Moderate to strong Id-3 immunoreactivity was present in the duct-like cancer cells (G and I, left panel). Moderate to strong Id-3 immunoreactivity was also present in the ductal cells of normal pancreatic tissue samples (H). Id-3 immunoreactivity was completely abolished by preabsorption with the Id-3 specific blocking peptide (I, right panel). A, D, and G constitute serial sections of a pancreatic cancer sample, revealing coexpression of the three Id proteins. Scale bars, 25 μm.

Cruz, CA) polyclonal antibodies were used for immunhistochemistry. These affinity-purified rabbit polyclonal antibodies specifically react with Id-1, Id-2, and Id-3, respectively, of human origin, as determined by Western blotting. Paraffin-embedded sections (4 μm) were subjected to immunostaining using the streptavidin-peroxidase technique. Where indicated, immunostaining for all three Id proteins was performed on serial sections. Endogenous peroxidase activity was blocked by incubation for 30 minutes with 0.3% hydrogen peroxide in methanol. Tissue sections were incubated for 15 minutes (23°C) with 10% normal goat serum and then incubated for 16 hours at 4°C with the indicated antibodies in PBS containing 1% bovine serum albumin. Bound antibodies were detected with biotinylated goat anti-rabbit IgG secondary antibodies and streptavidin-peroxidase complex, using diaminobenzidine tetrahydrochloride as the substrate. Sections were counterstained with Mayer's hematoxylin. Preabsorption with Id-1-, Id-2-, or Id-3-specific blocking peptides completely abolished immunoreactivity of the respective primary antibody. The immunohistochemical results were semiquantitatively analyzed as described previously.[29,30] The percentage of positive cancer cells was stratified into four groups: 0, no cancer cells exhibiting immunoreactivity; 1, <33% of the cancer cells exhibiting immunoreactivity; 2, 33 to 67% of the cancer cells exhibiting immunoreactivity; 3 >67% of the cancer cells exhibiting immunoreactivity. The intensity of the immunohistochemical signal was also stratified into four groups: 0, no immunoreactivity; 1, weak immunoreactivity; 2, moderate immunoreactivity; 3, strong immunoreactivity. Finally, the sum of the results of the cell score and the intensity score was calculated. Statistical analysis was performed with SigmaStat software. The rank sum test was used, and $P < 0.05$ was taken as the level of significance.

Results

Northern blot analysis of total RNA isolated from 12 normal pancreatic tissues and 16 pancreatic cancers revealed the presence of the 1.2-kb Id-1 transcript and the 1.6-kb Id2 mRNA transcript in 11 of the 12 normal pancreatic samples, and the 1.3-kb Id-3 mRNA transcript in all normal pancreatic samples (Figure 1A, 2). In the cancer tissues, Id-1 mRNA levels were elevated in 8 of 16 samples, Id-2 mRNA levels were elevated in 9 of these samples, and Id-3 mRNA levels were elevated in 6 of these samples (Figure 1A, 2). Concomitant overexpression of all three Id species was observed in 6 of the cancer samples (38%). In contrast, none of the Id mRNA species were overexpressed in CP by comparison with normal controls (Figure 1B, 2). Densitometric analysis of all of the autoradiograms indicated that there was a 6.5-fold increase ($P < 0.01$) in Id-1 mRNA levels, a fivefold increase ($P < 0.01$) in Id-2 mRNA levels, and a twofold increase ($P = 0.027$) in Id-3 mRNA levels in the pancreatic cancer samples in comparison to normal controls (Figure 2). In contrast, there was no statistically significant difference in the expression levels of Id-1, Id-2, and Id-3, in CP tissues in comparison to the corresponding levels in the normal pancreas (Figure 2).

Next, we assessed the expression of the three Id genes in 5 human pancreatic cancer cell lines by Northern and Western blot analyses. Id-1 mRNA was present at varying levels in all 5 cell lines (Figure 3). ASPC-1, CAPAN-1, MIA-PaCa-2, and PANC-1 expressed moderate to high levels of Id-1 mRNA, whereas COLO-357 cells

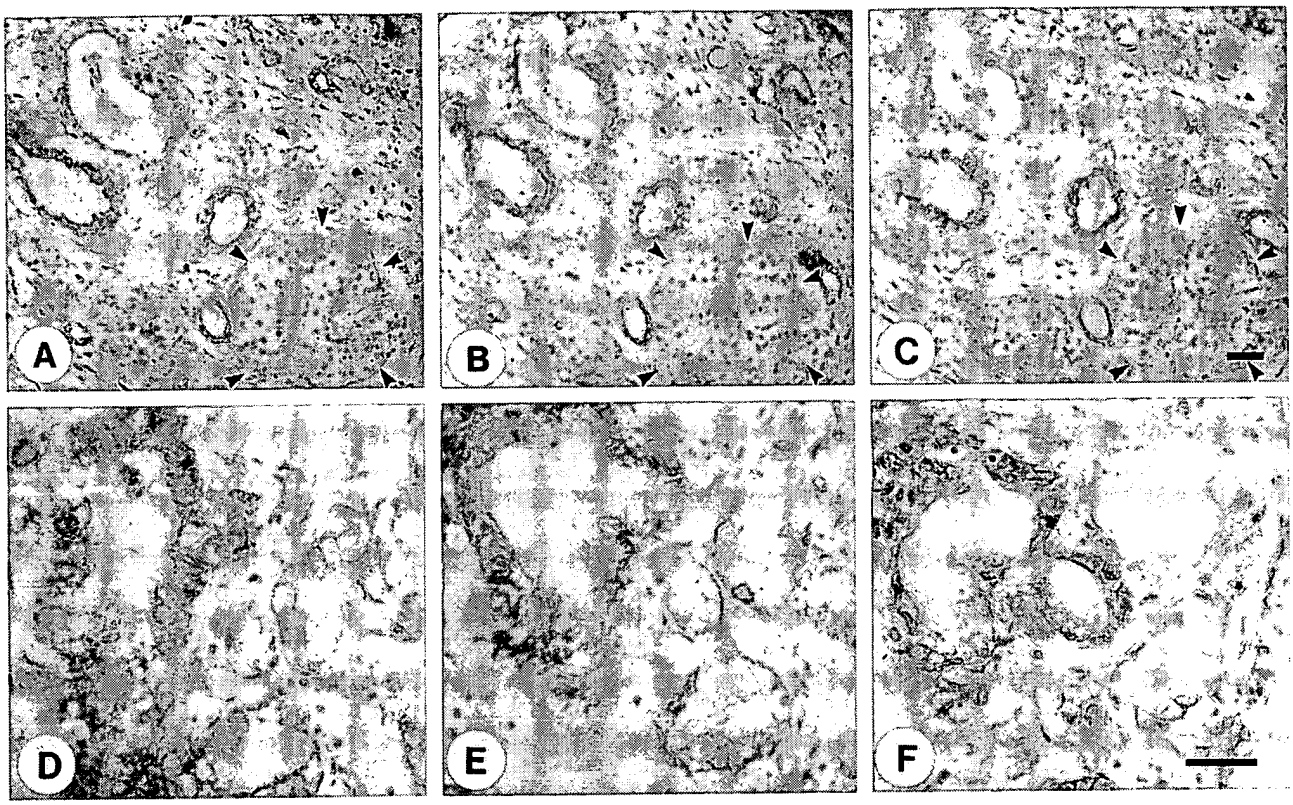


Figure 5. Immunohistochemistry of pancreatic cancer and dysplastic ducts in CP tissues. In the pancreatic cancer tissues (**A-C**) there was moderate to strong Id-1 (**A**), Id-2 (**B**), and Id-3 (**C**) immunoreactivity in the ductal cells in the areas adjacent to the cancer cells that exhibited CP-like alterations. Islet cells did not exhibit Id immunoreactivity (outlined by **solid arrowheads**). In the CP samples, moderate to strong Id-1 (**D**), Id-2 (**E**), and Id-3 (**F**) immunoreactivity was present in the cytoplasm of epithelial cells forming large dysplastic ducts. Scale bar, 25 μm.

expressed relatively low levels of this mRNA moiety. Western blotting with a highly specific anti-Id-1 antibody confirmed the presence of the approximately 14-kd Id-1 protein in the 4 cell lines that expressed high levels of Id-1 mRNA (Figure 3). Furthermore, the three cell lines with the highest Id-1 mRNA expression (CAPAN-1, MIA-PaCa-2, and PANC-1) also exhibited the highest Id-1 protein expression. Variable levels of the 1.6-kb Id-2 mRNA transcript were present in all 5 cell lines. In addition, a minor band of approximately 1.2 kb was visible in COLO-357 and MIA-PaCa-2 cells. Immunoblot analysis with a highly specific anti-Id-2 antibody revealed two bands of approximately 16 and 18 kd at relatively high levels in all of the cell lines with exception of PANC-1 cells, in which the 16-kd band was relatively faint (Figure 3). With the exception of MIA-PaCa-2 cells, there was a good correlation between Id-2 mRNA and protein levels (Figure 3). Id-3 mRNA was present at high levels in MIA-PaCa-2 cells, at moderate levels in COLO-357 cells, and at low levels in PANC-1 cells. Id-3 mRNA was not detectable in ASPC-1 and CAPAN-1 cells (Figure 3). Immunoblot analysis with a highly specific anti-Id-3 antibody revealed an approximately 14-kd band that was most abundant in MIA-PaCa-2 cells, and was also readily apparent in COLO-357 and PANC-1 cells. In contrast, only a faint Id-3 band was seen in ASPC-1 and CAPAN-1 cells. Thus, with the exception of PANC-1 cells, there was a good correlation between Id-3 mRNA and protein levels.

To determine the localization of Id-1, Id-2, and Id-3, immunostaining was carried out using the same highly specific anti-Id antibodies. In the pancreatic cancers, moderate to strong Id-1 immunoreactivity was present in the cancer cells in 9 of 10 randomly selected cancer samples. An example of moderate Id-1 immunoreactivity is shown in Figure 4A, and of strong immunoreactivity in Figure 4C (left panel). In contrast, in the normal pancreas, faint Id-1 immunoreactivity was present only in the ductal cells of pancreatic ducts (Figure 4B, arrowheads). Preabsorption with the Id-1-specific blocking peptide completely abolished the Id-1 immunoreactivity (Figure 4C, right panel). The cancer cells also exhibited strong Id-2 (Figure 4, D and F, left panel) and moderate to strong Id-3 immunoreactivity. An example of moderate Id-3 immunoreactivity is shown in Figure 4G, and of strong immunoreactivity in Figure 4I (left panel). In contrast, only faint Id-2 immunoreactivity was present in the ductal cells in the normal pancreas (Figure 4E), whereas Id-3 immunoreactivity in these cells was more variable and ranged from moderate to occasionally strong (Figure 4H). Islet cells and acinar cells were always devoid of Id immunoreactivity. Preabsorption of the respective antibody with the blocking peptides specific for Id-2 (Figure 4F, right panel) and Id-3 (Figure 4I, right panel) completely abolished immunoreactivity. Analysis of serial pancreatic cancer sections revealed that there was often colocalization of the

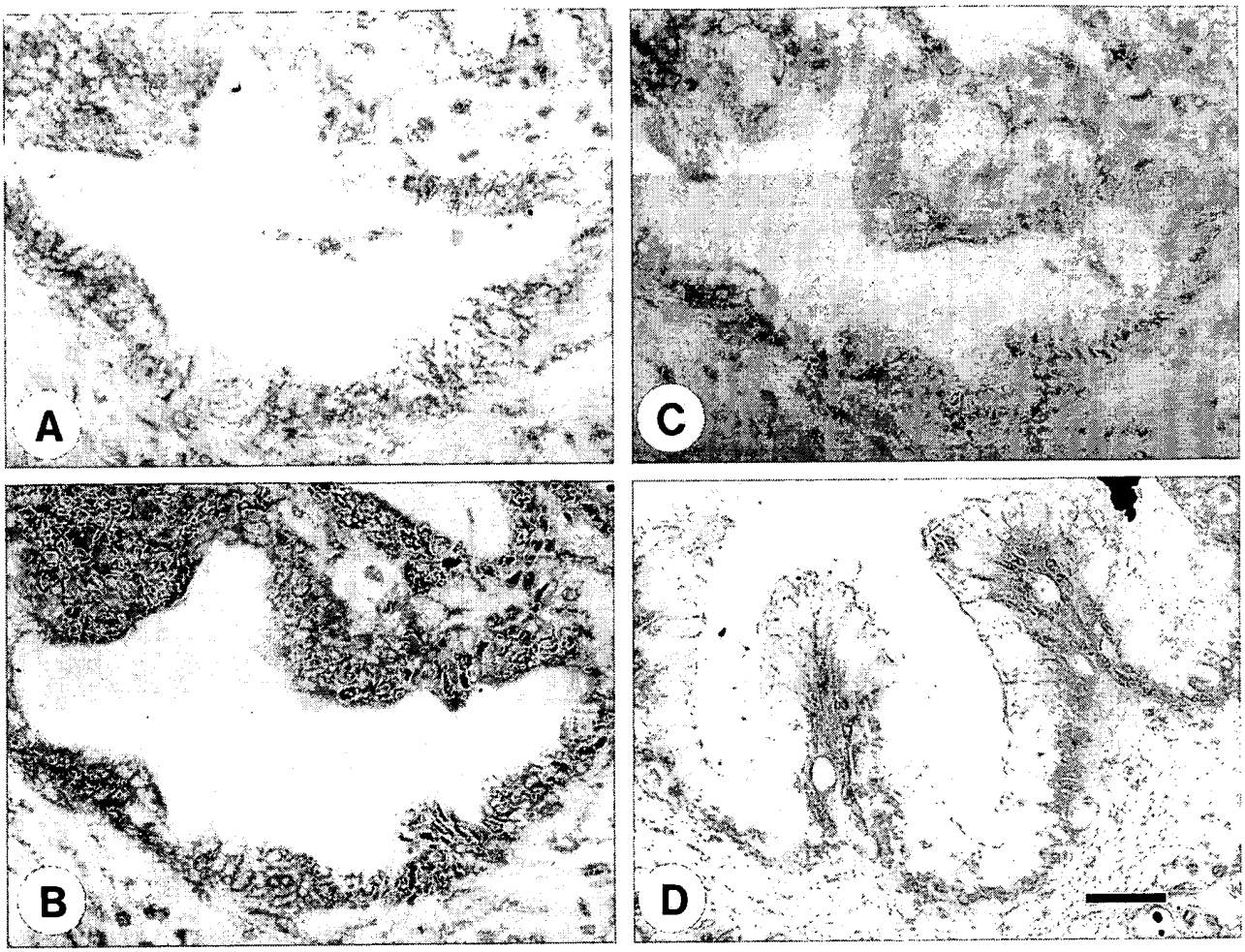


Figure 6. Immunohistochemistry of atypical papillary epithelium in CP tissues. Serial section analysis of some CP samples revealed the presence of large duct-like structures with atypical papillary epithelium. Mild to moderate Id-1 (A) and Id-2 (B) immunoreactivity and weak Id-3 (C) immunoreactivity was present in the cytoplasm of the cells forming these large ducts with papillary structures. Some CP samples also exhibited moderate Id-3 immunoreactivity in these cells (D). Scale bar, 25 μm.

three Id proteins. An example of serial sections from a pancreatic cancer tissue is shown in Figure 4, A, D, and G.

Id-1, Id-2, and Id-3 immunoreactivity was also present at moderate levels in the cytoplasm of ductal cells within CP-like areas adjacent to the cancer cells (Figure 5, A-C). As in the normal pancreas, islet cells (outlined by arrowheads) did not exhibit Id immunoreactivity. In 4 of 9 CP samples, there were foci of ductal cell dysplasia of relatively large interlobular ducts, all of which exhibited moderate to strong Id-1, Id-2, and Id-3 immunoreactivity (Figure 5, D-F). Five of 9 CP samples also contained foci of large ducts exhibiting atypical papillary epithelium. Serial section analysis of one of those CP samples revealed mild to moderate Id-1 and Id-2 immunoreactivity and weak Id-3 immunoreactivity in the cells of these atypical papillary ducts (Figure 6, A-C). In contrast, in some of these CP samples, moderate to strong Id-3 immunoreactivity was also observed (Figure 6D). However, most of the ductal cells forming the typical ductular structures of CP, such as large interlobular ducts and small proliferating ducts, exhibited generally only weak to occasionally moderate Id immunoreactivity (data not shown).

The immunohistochemical data for Id-1, Id-2, and Id-3 are summarized in Table 1. In the case of Id-1 and Id-2, the cancer cells as well as the dysplastic and atypical papillary ducts in CP exhibited a significantly higher score than the ductal cells in the normal pancreas. In contrast, due to the marked variability in Id-3 immunostaining in the normal pancreas, the differences between normal and cancer cells and normal and dysplastic cells did not achieve statistical significance.

Discussion

Id proteins constitute a family of HLH transcription factors that are important regulators of cellular differentiation and proliferation.[1,2] To date, four members of the human Id family have been identified.[1,10–12] Their expression is enhanced during cellular proliferation and in response to mitogenic stimuli,[19,31] and overexpression of Id genes inhibits differentiation and/or enhances proliferation in several different cell types.[15,32–34] The forced expression of Id-1 in mouse small intestinal epithelium results in

Table 1. Histological Scoring

		Id-1	Id-2	Id-3
Normal (n = 6)	Ductal cells	2.0 ± 0.4	2.3 ± 0.2	2.5 ± 0.9
Cancer (n = 10)	Cancer cells	4.5* ± 0.5	5.2§ ± 0.3	4.5 ± 0.6
CP (n = 9)	Typical CP lesions (n = 9)	2.7 ± 0.5	3.1 ± 0.6	3.4 ± 0.7
	Dysplastic ducts (n = 4)	5.3† ± 0.2	5.8‡ ± 0.2	5.3 ± 0.4
	Atypical papillary ducts (n = 5)	4.4‡ ± 0.2	5.2‡ ± 0.2	5.0 ± 0.4

Scoring of the histological specimens was performed as described in the Patients and Methods section. Values are the means ± SD of the number of samples indicated in parenthesis. P values are based on comparisons with the respective controls in the normal samples.
*, $P < 0.02$; †$P < 0.01$; ‡$P = 0.004$; §$P = 0.001$.

adenoma formation in these animals.[35] The growth-promoting effects of Id genes are thought to occur through several mechanisms. For example, Id-2 can bind to members of the pRB tumor suppressor family, thus blocking their growth-suppressing activity,[20,21] and Id-1 and Id-2 can antagonize the bHLH-mediated activation of known inhibitors of cell cycle progression such as the cyclin-dependent kinase inhibitor p21.[23]

In the present study, we determined by Northern blot analysis that a significant percentage of human pancreatic cancers expressed increased Id-1, Id-2, and Id-3 mRNA levels. Increased expression was most evident for Id-1 (6.5-fold) and Id-2 (fivefold). In contrast, Id-3 mRNA levels were only twofold increased in the cancer samples, partly because this mRNA was present at relatively high levels in the normal pancreas. Immunhistochemical analysis confirmed the presence of Id-1, Id-2, and Id-3 in the cancer cells within the tumor mass, whereas in the normal pancreas faint Id-1 and Id-2 immunoreactivity and moderate to occasionally strong Id-3 immunoreactivity was present in some ductal cells. Pancreatic acinar and islet cells in the normal pancreas were devoid of Id-1, Id-2, and Id-3 immunoreactivity. In the cancer samples, all three Id proteins often colocalized in the cancer cells. Coexpression of all three Id genes was also observed in cultured pancreatic cancer cell lines, which often exhibited a close correlation between Id mRNA and protein expression. However, in MIA-PaCa-2 there was a divergence of Id-2 mRNA and protein levels, and in PANC-1 cells, Id-3 mRNA levels did not correlate well with Id-3 protein expression. These observations suggest that in these cells, the half-life of either Id mRNA or Id protein may be altered by comparison with the other cell lines. Interestingly, Id-2 immunoblotting revealed two closely spaced bands of approximately 16 and 18 kd in 4 of 5 cell lines. In view of the fact that two possible initiation codons have been reported for the Id-2 gene,[36] our observation raises the possibility that the two Id-2-immunoreactive bands may represent separate translation products of the Id-2 gene.

Pancreatic cancers often harbor p53 tumor suppressor gene mutations[37] and exhibit alterations in apoptosis pathways. Thus, these cancers often exhibit increased expression of anti-apoptotic proteins such as Bcl-2[38] and abnormal resistance to Fas-ligand-mediated apoptosis.[39] It has been shown recently that forced constitutive expression of Id genes together with the expression of anti-apoptotic genes such as Bcl-2 or $BclX_L$ can result in malignant transformation of human fibroblasts,[11] raising the possibility that the enhanced Id expression in pancreatic cancers together with increased expression of anti-apoptotic genes may contribute to the malignant potential of pancreatic cancer cells *in vivo*.

In the CP tissues there was no significant increase in Id-1, Id-2, and Id-3 mRNA levels in comparison to the normal pancreas. Immunohistochemical analysis of pancreatic cancer samples revealed colocalization of weak to moderate Id-1, Id-2, and Id-3 immunoreactivity in proliferating ductal cells in the CP-like regions adjacent to the cancer cells, indicating that Id expression was not restricted to the cancer cells. Similarly, analysis of CP samples indicated weak Id-1, Id-2, and Id-3 immunoreactivity in the cells of small proliferating ducts and large ducts without dysplastic changes. In general, there was a correlation between weak immunoreactivity and low Id mRNA levels. However, in samples that harbored large ducts with papillary structures there was moderate Id immunoreactivity, and in the cells forming dysplastic ducts there was moderate to strong Id immunoreactivity. In these CP samples, Id mRNA levels were relatively higher than in the CP samples that were devoid of these histological changes. Overall, however, increased Id expression, most notably of Id-1 and Id-2, distinguished a subgroup of pancreatic cancers from CP (Table 1).

Epidemiological studies have shown that the risk of developing pancreatic cancer is increased up to 16-fold in patients with pre-existing CP in comparison to the general population.[40] The mechanisms that contribute to neoplastic transformation in CP are not known. Although there is no established tumor progression model for pancreatic cancer, such as the adenoma-carcinoma sequence of colorectal carcinoma,[41] it is generally accepted that K-*ras* and p16 mutations occur relatively early in pancreatic carcinogenesis, whereas p53 mutations occur late in this process.[37,41–43] Increased Id expression may contribute to malignant transformation of cultured cell lines *in vitro*[11] and has been linked to cell invasion in a murine mammary epithelial cell line.[44] In view of the current findings that Id-1, Id-2, and Id-3 are overexpressed in pancreatic cancer and in dysplastic/metaplastic ducts in CP, these observations raise the possibility that elevated levels of Id-1, Id-2, and, to a lesser extent, Id-3 may represent relatively early markers of pancreatic malignant transformation and may contribute to the pathobiology of pancreatic cancer.

References

1. Jan YN, Jan LY: HLH proteins, fly neurogenesis, and vertebrate myogenesis. Cell 1993, 75:827–830
2. Olson EN, Klein WH: bHLH factors in muscle development: dead lines and commitments, what to leave in and what to leave out. Genes Dev 1994, 8:1–8
3. Begley CG, Aplan PD, Denning SM, Haynes BF, Waldmann TA, Kirsch IR: The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. Proc Natl Acad Sci USA 1989, 86:10128–10132
4. Johnson JE, Birren SJ, Anderson DJ: Two rat homologues of Drosophila achaete-scute specifically expressed in neuronal precursors. Nature 1990, 346:858–861
5. Weintraub H: The MyoD family and myogenesis: redundancy, networks, and thresholds. Cell 1993, 75:1241–1244
6. Hu JS, Olson EN, Kingston RE: HEB, a helix-loop-helix protein related to E2A, and ITF2 that can modulate the DNA-binding ability of myogenic regulatory factors. Mol Cell Biol 1992, 12:1031–1042
7. Langlands K, Yin X, Anand G, Prochownik EV: Differential interactions of Id proteins with basic-helix-loop-helix transcription factors. J Biol Chem 1997, 272:19785–19793
8. Murre C, Bain G, van Dijk MA, Engel I, Furnari BA, Massari ME, Matthews JR, Quong MW, Rivera RR, Stuiver MH: Structure and function of helix-loop-helix proteins. Biochim Biophys Acta 1994, 1218:129–135
9. Murre C, McCaw PS, Vaessin H, Caudy M, Jan LY, Jan YN, Cabrera CV, Buskin JN, Hauschka SD, Lassar AB, Baltimore D: Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell 1989, 58: 537–544
10. Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H: The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell 1990, 61:49–59
11. Norton JD, Atherton GT: Coupling of cell growth control and apoptosis functions of Id proteins. Mol Cell Biol 1998, 18:2371–2381
12. Norton JD, Deed RW, Craggs G, Sablitzky F: Id helix-loop-helix proteins in cell growth and differentiation. Trends Cell Biol 1998, 8:58–65
13. Christy BA, Sanders LK, Lau LF, Copeland NG, Jenkins NA, Nathans D: An Id-related helix-loop-helix protein encoded by a growth factor-inducible gene. Proc Natl Acad Sci USA 1991, 88:1815–1819
14. Kawaguchi N, DeLuca HF, Noda M: Id gene expression and its suppression by 1,25-dihydroxyvitamin D3 in rat osteoblastic osteosarcoma cells. Proc Natl Acad Sci USA 1992, 89:4569–4572
15. Kreider BL, Benezra R, Rovera G, Kadesch T: Inhibition of myeloid differentiation by the helix-loop-helix protein Id. Science 1992, 255: 1700–1702
16. Le Jossic C, Ilyin GP, Loyer P, Glaise D, Cariou S, Guguen-Guillouzo C: Expression of helix-loop-helix factor Id-1 is dependent on the hepatocyte proliferation and differentiation status in rat liver and in primary culture. Cancer Res 1994, 54:6065–6068
17. Sun XH, Copeland NG, Jenkins NA, Baltimore D: Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. Mol Cell Biol 1991, 11:5603–5611
18. Wilson RB, Kiledjian M, Shen CP, Benezra R, Zwollo P, Dymecki SM, Desiderio SV, Kadesch T: Repression of immunoglobulin enhancers by the helix-loop-helix protein Id: implications for B-lymphoid-cell development. Mol Cell Biol 1991, 11:6185–6191
19. Hara E, Yamaguchi T, Nojima H, Ide T, Campisi J, Okayama H, Oda K: Id-related genes encoding helix-loop-helix proteins are required for G1 progression and are repressed in senescent human fibroblasts. J Biol Chem 1994, 269:2139–2145
20. Iavarone A, Garg P, Lasorella A, Hsu J, Israel MA: The helix-loop-helix protein Id-2 enhances cell proliferation and binds to the retinoblastoma protein. Genes Dev 1994, 8:1270–1284
21. Lasorella A, Iavarone A, Israel MA: Id2 specifically alters regulation of the cell cycle by tumor suppressor proteins. Mol Cell Biol 1996, 16:2570–2578
22. Peverali FA, Ramqvist T, Saffrich R, Pepperkok R, Barone MV, Philipson L: Regulation of G1 progression by E2A and Id helix-loop-helix proteins. EMBO J 1994, 13:4291–4301
23. Prabhu S, Ignatova A, Park ST, Sun XH: Regulation of the expression of cyclin-dependent kinase inhibitor p21 by E2A and Id proteins. Mol Cell Biol 1997, 17:5888–5896
24. Warshaw AL, Fernandez-del Castillo C: Pancreatic carcinoma. N Engl J Med 1992, 326:455–465
25. Korc M: Role of growth factors in pancreatic cancer. Surg Oncol Clin North Am 1998, 7:25–41
26. Kleeff J, Ishiwata T, Friess H, Büchler MW, Israel MA, Korc M: The helix-loop-helix protein Id2 is overexpressed in human pancreatic cancer. Cancer Res 1998, 58:3769–3772
27. Oertel JE, Heffes CS, Oertel YC: Pancreas. Diagnostic Surgical Pathology. Edited by SS Sternberg. New York, Raven Press, 1989, pp 1057–1093
28. Korc M, Chandrasekar B, Yamanaka Y, Friess H, Büchler MW, Beger HG: Overexpression of the epidermal growth factor receptor in human pancreatic cancer is associated with concomitant increase in the levels of epidermal growth factor and transforming growth factor α. J Clin Invest 1992, 90:1352–1360
29. Saeki T, Stromberg K, Qi CF, Gullick WJ, Tahara E, Normanno N, Ciardiello F, Kenney N, Johmson GR, Salomon DS: Differential immunohistochemical detection of amphiregulin and cripto in human normal colon and colorectal tumors. Cancer Res 1992, 52:3467–3473
30. Cantero D, Friess H, Deflorin J, Zimmermann A, Bründler MA, Riesle E, Korc M, Büchler MW: Enhanced expression of urokinase plasminogen activator and its receptor in pancreatic carcinoma. Br J Cancer 1997, 75:388–395
31. Desprez PY, Hara E, Bissell MJ, Campisi J: Suppression of mammary epithelial cell differentiation by the helix-loop-helix protein Id-1. Mol Cell Biol 1995, 15:3398–3404
32. Shoji W, Yamamoto T, Obinata M: The helix-loop-helix protein Id inhibits differentiation of murine erythroleukemia cells. J Biol Chem 1994, 269:5078–5084
33. Cross JC, Flannery ML, Blanar MA, Steingrimsson E, Jenkins NA, Copeland NG, Rutter WJ, Werb Z: Hxt encodes a basic helix-loop-helix transcription factor that regulates trophoblast cell development. Development 1995, 121:2513–2523
34. Sun XH: Constitutive expression of the Id1 gene impairs mouse B cell development. Cell 1994, 79:893–900
35. Wice BM, Gordon JI: Forced expression of Id-1 in the adult mouse small intestinal epithelium is associated with development of adenomas. J Biol Chem 1998, 273:25310–25319
36. Barone MV, Pepperkok R, Peverali FA, Philipson L: Id proteins control growth induction in mammalian cells. Proc Natl Acad Sci USA 1994, 91:4985–4988
37. Barton CM, Staddon SL, Hughes CM, Hall, PA, O'Sullivan C, Kloppel G, Theis B, Russell RC, Neoptolmos J, Williamson RCN, Lane DP, Lemoine NR: Abnormalities of the p53 tumour suppressor gene in human pancreatic cancer. Br J Cancer 1991, 64:1076–1082
38. Ohshio G, Suwa H, Imamura T, Yamaki K, Tanaka T, Hashimoto Y, Imamura M: An immunohistochemical study of bcl-2 and p53 protein expression in pancreatic carcinomas. Scand J Gastroenterol 1998, 33:535–539
39. Ungefroren H, Voss M, Jansen M, Roeder C, Henne-Bruns D, Kremer B, Kalthoff H: Human pancreatic adenocarcinomas express Fas and Fas ligand yet are resistant to Fas-mediated apoptosis. Cancer Res 1998, 58:1741–1749
40. Niederau C, Niederau MC, Heintges T, Lüthen R: Epidemiology: relation between chronic pancreaitis and pancreatic carcinoma. Cancer of the Pancreas. Edited by HG Beger, MW Büchler, MH Schoenberg. Ulm, Germany, Universitätsverlag Ulm GmbH, 1996, pp 6–9
41. Moskaluk CA, Kern SE: Molecular genetics of pancreatic carcinoma. Pancreatic Cancer: Pathogenesis, Diagnosis, and Treatment. Edited by HA Reber. Totowa, NJ, Humana Press, 1998, pp 3–20
42. Moskaluk CA, Hruban RH, Kern SE: p16 and K-ras gene mutations in the intraductal precursors of human pancreatic adenocarcinoma. Cancer Res 1997, 57:2140–2143
43. Tada M, Ohashi M, Shiratori Y, Okudaira T, Komatsu Y, Kawabe T, Yoshida H, Machinami R, Kishi K, Omata M: Analysis of K-ras gene mutation in hyperplastic duct cells of the pancreas without pancreatic disease. Gastroenterology 1996, 110:227–231
44. Desprez PY, Lin CQ, Thomasset N, Sympson CJ, Bissell MJ, Campisi J: A novel pathway for mammary epithelial cell invasion induced by the helix-loop-helix protein Id-1. Mol Cell Biol 1998, 18:4577–4588

MOLECULAR AND CELLULAR BIOLOGY, Nov. 1999, p. 7357–7368 Vol. 19, No. 11
0270-7306/99/$04.00+0

A Sampling of the Yeast Proteome

B. FUTCHER,[1]* G. I. LATTER,[1] P. MONARDO,[1] C. S. McLAUGHLIN,[2] AND J. I. GARRELS[3]

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724[1]; Department of Biological Chemistry, University of California, Irvine, California 92717[2]; and Proteome, Inc., Beverly, Massachusetts 01915[3]

Received 15 June 1999/Returned for modification 16 July 1999/Accepted 28 July 1999

In this study, we examined yeast proteins by two-dimensional (2D) gel electrophoresis and gathered quantitative information from about 1,400 spots. We found that there is an enormous range of protein abundance and, for identified spots, a good correlation between protein abundance, mRNA abundance, and codon bias. For each molecule of well-translated mRNA, there were about 4,000 molecules of protein. The relative abundance of proteins was measured in glucose and ethanol media. Protein turnover was examined and found to be insignificant for abundant proteins. Some phosphoproteins were identified. The behavior of proteins in differential centrifugation experiments was examined. Such experiments with 2D gels can give a global view of the yeast proteome.

The sequence of the yeast genome has been determined (9). More recently, the number of mRNA molecules for each expressed gene has been measured (27, 30). The next logical level of analysis is that of the expressed set of proteins. We have begun to analyze the yeast proteome by using two-dimensional (2D) gels.

2D gel electrophoresis separates proteins according to isoelectric point in one dimension and molecular weight in the other dimension (21), allowing resolution of thousands of proteins on a single gel. Although modern imaging and computing techniques can extract quantitative data for each of the spots in a 2D gel, there are only a few cases in which quantitative data have been gathered from 2D gels. 2D gel electrophoresis is almost unique in its ability to examine biological responses over thousands of proteins simultaneously and should therefore allow us a relatively comprehensive view of cellular metabolism.

We and others have worked toward assembling a yeast protein database consisting of a collection of identified spots in 2D gels and of data on each of these spots under various conditions (2, 7, 8, 10, 23, 25). These data could then be used in analyzing a protein or a metabolic process. *Saccharomyces cerevisiae* is a good organism for this approach since it has a well-understood physiology as well as a large number of mutants, and its genome has been sequenced. Given the sequence and the relative lack of introns in *S. cerevisiae*, it is easy to predict the sequence of the primary protein product of most genes. This aids tremendously in identifying these proteins on 2D gels.

There are three pillars on which such a database rests: (i) visualization of many protein spots simultaneously, (ii) quantification of the protein in each spot, and (iii) identification of the gene product for each spot. Our first efforts at visualization and identification for *S. cerevisiae* have been described elsewhere (7, 8). Here we describe quantitative data for these proteins under a variety of experimental conditions.

MATERIALS AND METHODS

Strains and media. *S. cerevisiae* W303 (*MATa ade2-1 his3-11,15 leu2-3, 112 trp1-1 ura3-1 can1-100*) was used (26). −Met YNB (yeast nitrogen base) medium was 1.7 g of YNB (Difco) per liter, 5 g of ammonium sulfate per liter, and adenine, uracil, and all amino acids except methionine; −Met −Cys YNB medium was the same but without methionine or cysteine. Medium was supplemented with 2% glucose (for most experiments) or with 2% ethanol (for ethanol experiments). Low-phosphate YEPD was described by Warner (28).

Isotopic labeling of yeast and preparation of cell extracts. Yeast strains were labeled and proteins were extracted as described by Garrels et al. (7, 8). Briefly, cells were grown to 5×10^6 cells per ml. at 30°C; 1 ml of culture was transferred to a fresh tube, and 0.3 mCi of [^{35}S]methionine (e.g., Express protein labeling mix; New England Nuclear) was added to this 1-ml culture. The cells were incubated for a further 10 to 15 min and then transferred to a 1.5-ml microcentrifuge tube, chilled on ice, and harvested by centrifugation. The supernatant was removed, and the cell pellet was resuspended in 100 μl of lysis buffer (20 mM Tris-HCl [pH 7.6], 10 mM NaF, 10 mM sodium pyrophosphate, 0.5 mM EDTA, 0.1% deoxycholate; just before use, phenylmethylsulfonyl fluoride was added to 1 mM, leupeptin was added to 1 μg/ml, pepstatin was added to 1 μg/ml, tosylsulfonyl phenylalanyl chloromethyl ketone was added to 10 μg/ml, and soybean trypsin inhibitor was added to 10 μg/ml).

The resuspended cells were transferred to a screw-cap 1.5-ml polypropylene tube containing 0.28 g of glass beads (0.5-mm diameter; Biospec Products) or 0.40 g of zirconia beads (0.5-mm diameter; Biospec Products). After the cap was secured, the tube was inserted into a MiniBeadbeater 8 (Biospec Products) and shaken at medium high speed at 4°C for 1 min. Breakage was typically 75%. Tubes were then spun in a microcentrifuge for 10 s at 5,000 × *g* at 4°C.

With a very fine pipette tip, liquid was withdrawn from the beads and transferred to a prechilled 1.5-ml tube containing 7 μl of DNase I (0.5 mg/ml; Cooper product no. 6330)–RNase A (0.25 mg/ml; Cooper product no. 5679)–Mg (50 mM $MgCl_2$) mix. Typically 70 μl of liquid was recovered. The mixture was incubated on ice for 10 min to allow the RNase and DNase to work.

Next, 75 μl of 2× dSDS (2× dSDS is 0.6% sodium dodecyl sulfate [SDS], 2% mercaptoethanol, and 0.1 M Tris-HCl [pH 8]) was added. The tube was plunged into boiling water, incubated for 1 min, and then plunged into ice. After cooling, the tube was centrifuged at 4°C for 3 min at 14,000 × *g*. The supernatant was transferred to a fresh tube and frozen at −70°C. About 5 μl of this supernatant was used for each 2D gel.

2D polyacrylamide gels. 2D gels were made and run as described elsewhere (6–8).

Image analysis of the gels. The Quest II software system was used for quantitative image analysis (20, 22). Two techniques were used to collect quantitative data for analysis by Quest II software. First, before the advent of phosphorimagers, gels were dried and fluorographed. Each gel was exposed to film for three different times (typically 1 day, 2 weeks, and 6 weeks) to increase the dynamic range of the data. The films were scanned along with calibration strips to relate film optical density to disintegrations per minute in the gels and analyzed by the software to obtain a linear relationship between disintegrations per minute in the spots and optical densities of the film images. The quantitative data are expressed as parts per million of the total cellular protein. This value is calculated from the disintegrations per minute of the sample loaded onto the gel and by comparing the film density of each data spot with density of the film over the calibration strips of known radioactivity exposed to the same film. This yields the disintegrations per minute per millimeter for each spot on the gel and thence its parts-per-minute value.

After the advent of phosphorimaging, gels bearing ^{35}S-labeled proteins were exposed to phosphorimager screens and scanned by a Fuji phosphorimager, typically for two exposures per gel. Calibration strips of known radioactivity were exposed simultaneously. Scan data from the phosphorimager was assimilated by Quest II software, and quantitative data were recorded for the spots on the gels.

* Corresponding author. Mailing address: Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724. Phone: (516) 367-8828. Fax: (516) 367–8369. E-mail: futcher@cshl.org.

Measurements of protein turnover. Cells in exponential phase were pulse-labeled with [^{35}S]methionine, excess cold Met and Cys were added, and samples of equal volume were taken from the culture at intervals up to 90 min (in one experiment) or up to 160 min (in a second experiment). Incorporation of ^{35}S into protein was essentially 100% by the first sample (10 min). Extracts were made, and equal fractions of the samples were loaded on 2D gels (i.e., the different samples had different amounts of protein but equal amounts of ^{35}S). Spots were quantitated with a phosphorimaging and Quest software.

The software was queried for spots whose radioactivity decreased through the time course. The algorithm examined all data points for all spots, drew a best-fit line through the data points, and looked for spots where this line had a statistically significant negative slope. In one of the experiments, there was one such spot. To the eye, this was a minor, unidentified spot seen only in the first two samples (10 and 20 min). In the other experiment, the Quest software found no spots meeting the criteria. Therefore, we concluded that none of the identified spots (and all but one of the visible spots) represented proteins with long half-lives.

Centrifugal fractionation. Cells were labeled, harvested, and broken with glass beads by the standard method described above except that no detergent (i.e., no deoxycholate) was present in the lysis buffer. The crude lysate was cleared of unbroken cells and large debris by centrifugation at 300 × *g* for 30 s. The supernatant of this centrifugation was then spun at 16,000 × *g* for 10 min to give the pellet used for Fig. 6B. The supernatant of the 16,000 × *g*, 10-min spin was then spun at 100,000 × *g* for 30 min to give the supernatant used for Fig. 6A.

Protein abundance calculations. A haploid yeast cell contains about 4×10^{-12} g of protein (1, 15). Assuming a mean protein mass of 50 kDa, there are about 50×10^6 molecules of protein per cell. There are about 1.8 methionines per 10 kDa of protein mass, which implies 4.5×10^8 molecules of methionine per cell (neglecting the small pool of free Met). We measured (i) the counts per minute in each spot on the 2D gels, (ii) the total number of counts on each gel (by integrating counts over the entire gel), and (iii) the total number of counts loaded on the gel (by scintillation counting of the original sample). Thus, we know what fraction of the total incorporated radioactivity is present in each spot. After correcting for the methionine (and cysteine [see below]) content of each protein, we calculated an absolute number of protein molecules based on the fraction of radioactivity in each spot and on 50×10^6 total molecules per cell.

The labeling mixture used contained about one-fifth as much radioactive cysteine as radioactive methionine. Therefore, the number of cysteine molecules per protein was also taken into account in calculating the number of molecules of protein, but Cys molecules were weighted one-fifth as heavily as Met molecules.

mRNA abundance calculations. For estimation of mRNA abundance, we used SAGE (serial analysis of gene expression) data (27) and Affymetrix chip hybridization data (29a, 30). The mRNA column in Table 1 shows mRNA abundance calculated from SAGE data alone. However, the SAGE data came from cells growing in YEPD medium, whereas our protein measurements were from cells growing in YNB medium. In addition, SAGE data for low-abundance mRNAs suffers from statistical variation. Therefore, we also used chip hybridization data (29a, 30) for mRNA from cells grown in YNB. These hybridization data also had disadvantages. First, the amounts of high-abundance mRNAs were systematically underestimated, probably because of saturation in the hybridizations, which used 10 μg of cRNA. For example, the abundance of *ADH1* mRNA was 197 copies per cell by SAGE but only 32 copies per cell by hybridization, and the abundance of *ENO2* mRNA was 248 copies per cell by SAGE but only 41 by hybridization. When the amount of cRNA used in the hybridization was reduced to 1 μg, the apparent amounts of mRNA were similar to the amounts determined by SAGE (29a, 29b). However, experiments using 1 μg of cRNA have been done for only some genes (29a). Because amounts of mRNA were normalized to 15,000 per cell, and because the amounts of abundant mRNAs were underestimated, there is a 2.2-fold overestimate of the abundance of nonabundant mRNAs. We calculated this factor of 2.2 by adding together the number of mRNA molecules from a large number of genes expressed at a low level for both SAGE data and hybridization data. The sum for the same genes from hybridization data is 2.2-fold greater than that from SAGE data.

To take into account these difficulties, we compiled a list of "adjusted" mRNA abundance as follows. For all high-abundance mRNAs of our identified proteins, we used SAGE data. For all of these particular mRNAs, chip hybridization suggested that mRNA abundance was the same in YEPD and YNB media. For medium-abundance mRNAs, SAGE data were used, but when hybridization data showed a significant difference between YEPD and YNB, then the SAGE data were adjusted by the appropriate factor. Finally, for low-abundance mRNAs, we used data from chip hybridizations from YNB medium but divided by 2.2 to normalize to the SAGE results. These calculations were completed without reference to protein abundance.

CAI. The codon adaptation index (CAI) was taken from the yeast proteome database (YPD) (13), for which calculations were made according to Sharp and Li (24). Briefly, the index uses a reference set of highly expressed genes to assign a value to each codon, and then a score for a gene is calculated from the frequency of use of the various codons in that gene (24).

Statistical analysis. The JMP program was used with the aid of T. Tully. The JMP program showed that neither mRNA nor protein abundances were normally distributed; therefore, Spearman rank correlation coefficients (r_s) were calculated. The mRNA (adjusted and unadjusted) and protein data were also transformed so that Pearson product-moment correlation coefficients (r_p) could be calculated. First, this was done by a Box-Cox transformation of log-transformed data. This transformation produced normal distributions, and an r_p of 0.76 was achieved. However, because the Box-Cox transformation is complex, we also did a simpler logarithmic transformation. This produced a normal distribution for the protein data. However, the distribution for the mRNA and adjusted mRNA data was close to, but not quite, normal. Nevertheless, we calculated the r_p and found that it was 0.76, identical to the coefficient from the Box-Cox transformed data. We therefore believe that this correlation coefficient is not misleading, despite the fact that the log(mRNA) distribution is not quite normal.

RESULTS

Visualization of 1,400 spots on three gel systems. Yeast proteins have isoelectric points ranging from 3.1 to 12.8, and masses ranging from less than 10 kDa to 470 kDa. It is difficult to examine all proteins on a single kind of gel, because a gel with the needed range in pI and mass would give poor resolution of the thousands of spots in the central region of the gel. Therefore, we have used three gel systems: (i) pH "4 to 8" with 10% polyacrylamide; (ii) pH "3 to 10" with 10% polyacrylamide; and (iii) nonequilibrium with 15% polyacrylamide (7, 8). Each gel system allows good resolution of a subset of yeast proteins.

Figure 1 shows a pH 4–8, 10% polyacrylamide gel. The pH at the basic end of the isoelectric focusing gel cannot be maintained throughout focusing, and so the proteins resolved on such gels have isoelectric points between pH 4 and pH 6.7. For these pH 4–8 gels, we see 600 to 900 spots on the best gels after multiple exposures.

The pH 3–10 gels (not shown) extend the pI range somewhat beyond pH 7.5, allowing detection of several hundred additional spots. Finally, we use nonequilibrium gels with 15% acrylamide in the second dimension. These allow visualization of about 100 very basic proteins and about 170 small proteins (less than 20 kDa). In total, using all three gel systems, about 1,400 spots can be seen. These represent about 1,200 different proteins, which is about one-quarter to one-third of the proteins expressed under these conditions (27, 30). Here, we focus on the proteins seen on the pH 4–8 gels.

Although nearly all expressed proteins are present on these gels, the number seen is limited by a problem we call coverage. Since there are thousands of proteins on each gel, many proteins comigrate or nearly comigrate. When two proteins are resolved, but are close together, and one protein spot is much more intense than the other, a problem arises in visualizing the weaker spot: at long exposures when the weak signal is strong enough for detection, the signal from the strong spot spreads and covers the signal from the weaker spot. Thus, weak spots can be seen only when they are well separated from strong spots.

For a given gel, the number of detectable spots initially rises with exposure time. However, beyond an optimal exposure, the number of distinguishable spots begins to decrease, because signals from strong spots cover signals from nearby weak spots. At long exposures, the whole autoradiogram turns black. Thus, there is an optimum exposure yielding the maximum number of spots, and at this exposure the weakest spots are not seen.

Largely because of the problem of coverage, the proteins seen are strongly biased toward abundant proteins. All identified proteins have a CAI of 0.18 or more, and we have identified no transcription factors or protein kinases, which are nonabundant proteins. Thus, this technology is useful for examining protein synthesis, amino acid metabolism, and glycolysis but not for examining transcription, DNA replication, or the cell cycle.

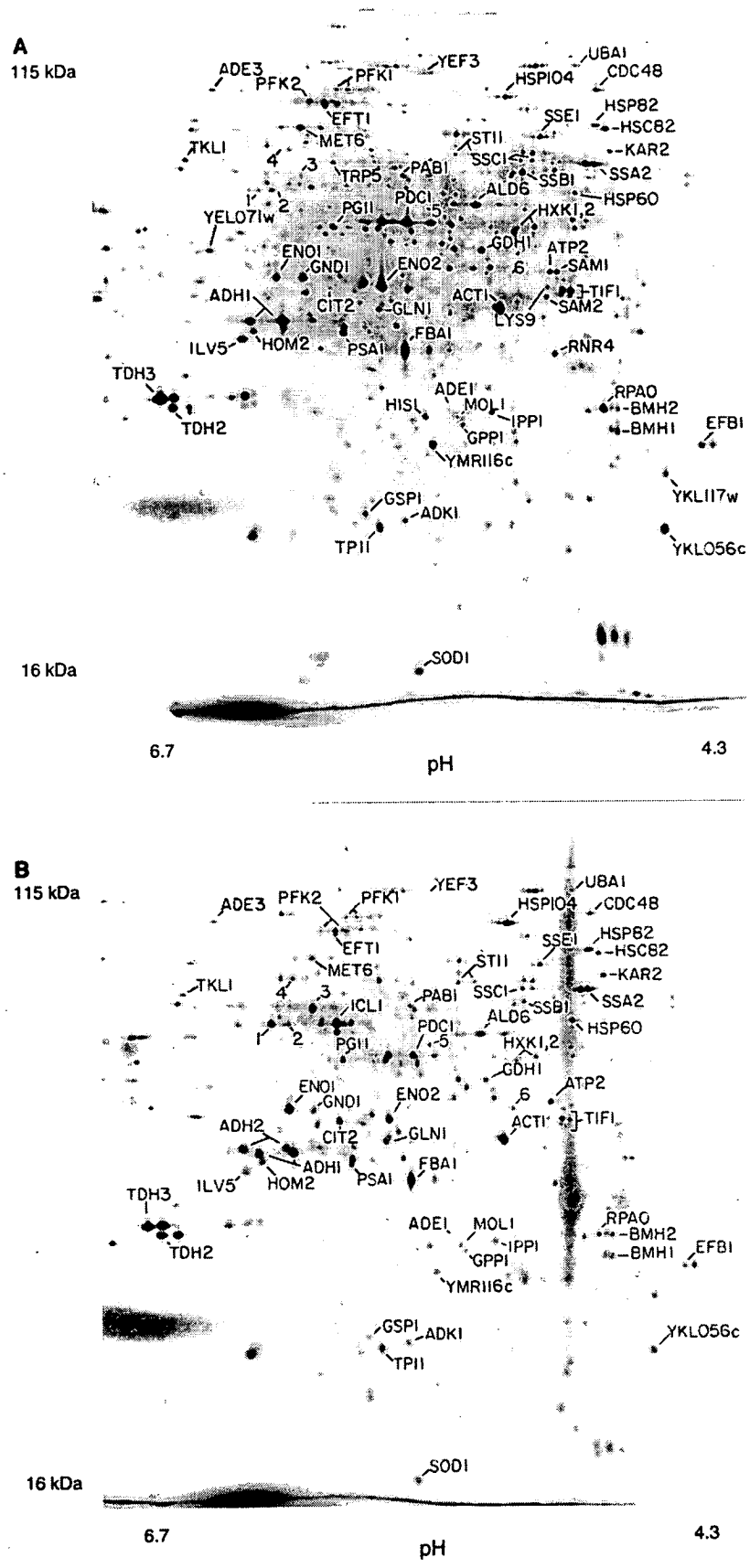


FIG. 1. 2D gels. The horizontal axis is the isoelectric focusing dimension, which stretches from pH 6.7 (left) to pH 4.3 (right). The vertical axis is the polyacrylamide gel dimension, which stretches from about 15 kDa (bottom) to at least 130 kDa (top). For panel A, extract was made from cells in log phase in glucose; for panel B, cells were grown in ethanol. The spots labeled 1 through 6 are unidentified proteins highly induced in ethanol.

Spot identification. The identification of various spots has been described elsewhere (7, 8). At present, 169 different spots representing 148 proteins have been identified. Many of these spots have been independently identified (2, 10, 23, 25). The main methods used in spot identification have been analysis of amino acid composition, gene overexpression, peptide sequencing, and mass spectrometry.

Pulse-chase experiments and protein turnover. Pulse-chase experiments were done to measure protein half-lives (Materials and Methods). Cells were labeled with [^{35}S]methionine for 10 min, and then an excess of unlabeled methionine was added. Samples were taken at 0, 10, 20, 30, 60, and 90 min after the beginning of the chase. Equal amounts of ^{35}S were loaded from each sample; 2D gels were run, and spots were quantitated. Surprisingly, almost every spot was nearly constant in amount of radioactivity over the entire time course (not shown). A few spots shifted from one position to another because of posttranslational modifications (e.g., phosphorylation of Rpa0 and Efb1). Thus, the proteins being visualized are all or nearly all very stable proteins, with half-lives of more than 90 min. Gygi et al. (10) have come to a similar conclusion by using the N-end rule to predict protein half-lives. This result does not imply that all yeast proteins are stable. The proteins being visualized are abundant proteins; this is partly because they are stable proteins.

Protein quantitation. Because all of the proteins seen had effectively the same half-life, the abundance of each protein was directly proportional to the amount of radioactivity incorporated during labeling. Thus, after taking into account the total number of protein molecules per cell, the average content of methionine and cysteine, and the methionine and cysteine content of each identified protein, we could calculate the abundance of each identified protein (Tables 1 and 2; Materials and Methods). About 1,000 unidentified proteins were also quantified, assuming an average content of Met and Cys.

Many proteins give multiple spots (7, 8). The contribution from each spot was summed to give the total protein amount. However, many proteins probably have minor spots that we are not aware of, causing the amount of protein to be underestimated.

When the proteins on a pH 4–8 gel were ordered by abundance, the most abundant protein had 8,904 ppm, the 10th most abundant had 2,842 ppm, the 100th most abundant had 314 ppm, the 500th most abundant had 57 ppm, and the 1,000th most abundant (visualized at greater than optimum exposure) had 23 ppm. Thus, there is more than a 300-fold range in abundance among the visualized proteins. The most abundant 10 proteins account for about 25% of the total protein on the pH 4–8 gel, the most abundant 60 proteins account for 50%, and the most abundant 500 proteins account for 80%. Since it seems likely that the pH 4–8 gels give a representative sampling of all proteins, we estimate that half of the total cellular protein is accounted for by fewer than 100 different gene products, principally glycolytic enzymes and proteins involved in protein synthesis.

Correlation of protein abundance with mRNA abundance. Estimates of mRNA abundance for each gene have been made by SAGE (27) and by hybridization of cRNA to oligonucleotide arrays (30). These two methods give broadly similar results, yet each method has strengths and weaknesses (Materials and Methods). Table 1 lists the number of molecules of mRNA per cell for each gene studied. One measurement (mRNA) uses data from SAGE analysis alone (27); a second incorporates data from both SAGE and hybridization (30) (adjusted mRNA) (Table 1; Materials and Methods). We correlated protein abundance with mRNA abundance (Fig. 2). For adjusted mRNA versus protein, the Spearman rank correlation coefficient, r_s, was 0.74 ($P < 0.0001$), and the Pearson correlation coefficient, r_p, on log transformed data (Materials and Methods) was 0.76 ($P < 0.00001$). We obtained similar correlations for mRNA versus protein and also for other data transformations (Materials and Methods). Thus, several statistical methods show a strong and significant correlation between mRNA abundance and protein abundance. Of course, the correlation is far from perfect; for mRNAs of a given abundance, there is at least a 10-fold range of protein abundance (Fig. 2). Some of this scatter is probably due to posttranscriptional regulation, and some is due to errors in the mRNA or protein data. For example, the protein Yef3 runs poorly on our gels, giving multiple smeared spots. Its abundance has probably been underestimated, partly explaining the low protein/mRNA ratio of Yef3. It is the most extreme outlier in Fig. 2.

These data on mRNA (27, 30) and protein abundance (Table 1) suggest that for each mRNA molecule, there are on average 4,000 molecules of the cognate protein. For instance, for Act1 (actin) there are about 54 molecules of mRNA per cell and about 205,000 molecules of protein. Assuming an mRNA half-life of 30 min (12) and a cell doubling time of 120 min, this suggests that an individual molecule of mRNA might be translated roughly 1,000 times. These calculations are limited to mRNAs for abundant proteins, which are likely to be the mRNAs that are translated best.

A full complement of cell protein is synthesized in about 120 min under these conditions. Thus, 4,000 molecules of protein per molecule of mRNA implies that translation initiates on an mRNA about once every 2 s. This is a remarkably high rate; it implies that if an average mRNA bears 10 ribosomes engaged in translation, then each ribosome completes translation in 20 s; if an average protein has 450 residues; this in turn implies translation of over 20 amino acids per s, a rate considerably higher than estimated for mammalians (3 to 8 amino acids per s) (18). These estimates depend on the amount of mRNA per cell (11, 27).

The large number of protein molecules that can be made from a single mRNA raises the issue of how abundance is controlled for less abundant proteins. Many nonabundant proteins may be unstable, and this would reduce the protein/mRNA ratio. In addition, many nonabundant proteins may be translated at suboptimal rates. We have found that mRNAs for nonabundant proteins usually have suboptimal contexts for translational initiation. For example, there are over 600 yeast genes which probably have short open reading frames in the mRNA upstream of the main open reading frame (17a). These may be devices for reducing the amount of protein made from a molecule of mRNA.

Correlation of codon bias with protein abundance. The mRNAs for highly expressed proteins preferentially use some codons rather than others specifying the same amino acid (14). This preference is called codon bias. The codons preferred are those for which the tRNAs are present in the greatest amounts. Use of these codons may make translation faster or more efficient and may decrease misincorporation. These effects are most important for the cell for abundant proteins, and so codon bias is most extreme for abundant proteins. The effect can be dramatic—highly biased mRNAs may use only 25 of the 61 codons.

We asked whether the correlation of codon bias with abundance continues for medium-abundance proteins. There are various mathematical expressions quantifying codon bias; here, we have used the CAI (24) (Materials and Methods) because it gives a result between 0 and 1. The r_s for CAI versus protein abundance is 0.80 ($P < 0.0001$), similar to the mRNA-protein

TABLE 1. Quantitative data[a]

Function	Name	CAI	mRNA	Adjusted mRNA	Protein (Glu) (10^3)	Protein (Eth) (10^3)	E/G ratio
Carbohydrate metabolism	Adh1	0.810	197	197	1,230	972	0.79
	Adh2	0.504	0		0	963	>20
	Cit2	0.185	1	2.8	23	288	12
	Eno1	0.870	No *Nla*		410	974	2.4
	Eno2	0.892	248	248	650	215	0.33
	Fba1	0.868	179	179	640	608	0.95
	Hxk1,2	0.500	13	10.5	62	46	
	Icl1	0.251	0		0	671	>20
	Pdb1	0.342	5	5	41	33	
	Pdc1	0.903	226	226	280	205	0.73
	Pfk1	0.465	5	5	75	53	0.71
	Pgi1	0.681	14	14	160	120	0.75
	Pyc1	0.260	1	0.7	37	34	
	Tal1	0.579	5	5	110	35	
	Tdh2	0.904	63	63	430	876	NR
	Tdh3	0.924	460	460	1,670	1,927	NR
	Tpi1	0.817	No *Nla*		No Met	No Met	
Protein synthesis	Efb1	0.762	33	16.5	358	362	
	Eft1,2	0.801	26	26	99	54	0.55
	Prt1	0.303	4	0.7	12	6	
	Rpa0	0.793	246	246	277	100	0.36
	Tif1,2	0.752	29	29	233	106	0.46
	Yef3	0.777	36	36	14	ND	
Heat shock	Hsc82	0.581	2	2.9	112	75	0.67
	Hsp60	0.381	9	2.3	35	82	2.3
	Hsp82	0.517	2	1.3	52	135	2.6
	Hsp104	0.304	7	7	70	161	2.3
	Kar2	0.439	5	10.1	43	102	2.4
	Ssa1	0.709	2	4.3	303	421	1.4
	Ssa2	0.802	10	5	213	324	1.5
	Ssb1,2	0.850	50	50	270	85	
	Ssc1	0.521	2	2.6	68	80	1.2
	Sse1	0.521	8	8	96	48	
	Sti1	0.247	1	1.1	25	44	1.7
Amino acid synthesis	Ade1	0.229	4	4	14	27	
	Ade3	0.276	2	1.7	12	9	
	Ade5,7	0.257	2	1.4	14	4	
	Arg4	0.229	1	8.1	41	41	
	Gdh1	0.585	10	27	148	55	
	Gln1	0.524	11	11	77	104	1.3
	His4	0.267	3	3	15	23	1.5
	Ilv5	0.801	6	6	152	109	0.7
	Lys9	0.332	4	4	32	17	0.52
	Met6	0.657	No *Nla*	22	190	80	0.42
	Pro2	0.248	3	3	30	12	
	Ser1	0.258	2	1.2	15	8	
	Trp5	0.319	5	5	28	12	
Miscellaneous	Act1	0.710	54	54	205	164	0.78
	Adk1	0.531	No *Nla*		47	43	
	Ald6	0.520	3	3	181	159	
	Atp2	0.424	1	4.1	76	109	1.4
	Bmh1	0.322	46	46	191	137	0.72
	Bmh2	0.384	1	1.4	134	147	
	Cdc48	0.306	2	2.4	32	26	
	Cdc60	0.299	2	0.86	6	2	
	Erg20	0.373	5	5	92	39	
	Gpp1	0.603	16	5	234	158	
	Gsp1	0.621	3	3	115	39	0.34
	Ipp1	0.620	4	4	254	147	0.58
	Lcb1	0.173	0.3	0.8	19	40	
	Mol1	0.423	0	0.45	20	16	
	Pab1	0.488	3	3	41	19	0.47
	Psa1	0.600	15	15	148	56	
	Rnr4	0.497	6	6	44	37	
	Sam1	0.494	5	5	59	21	
	Sam2	0.497	3	15	63	20	
	Sod1	0.376	36	36	631	618	
	Uba1	0.212	2	2	14	20	
	YKL056	0.731	62	62	253	112	0.44
	YLR109	0.549	21	21	930		
	YMR116	0.777	41	41	184	40	0.20

[a] CAI, a measure of codon bias, is taken from the YPD. mRNA, number of mRNA molecules per cell from SAGE data (27); adjusted mRNA, number of mRNA molecules per cell based on both SAGE and chip hybridization (30) (see Materials and Methods); Protein (Glu), number of molecules of protein per cell in YNB-glucose; Protein (Eth), number of molecules of protein per cell in YNB-ethanol; E/G ratio, ratio of protein abundance in ethanol to glucose. The E/G ratio is not given if it was close to 1 or if it was not repeatable (NR) in multiple gels. Some gene products (e.g., Tif1 and Tif2 [Tif1,2]) were difficult to distinguish on either a protein or an mRNA basis; these are pooled. No *Nla*, there was no suitable *Nla*III site in the 3′ region of the gene, and so there are no SAGE mRNA data; No Met, the mature gene product contains no methionines, and so there are no reliable protein data.

TABLE 2. Functions of proteins listed in Table 1

Name[a]	YPD title lines[b]
Adh1	Alcohol dehydrogenase I; cytoplasmic isozyme reducing acetaldehyde to ethanol, regenerating NAD^+
Adh2	Alcohol dehydrogenase II; oxidizes ethanol to acetaldehyde, glucose repressed
Cit2	Citrate synthase, peroxisomal (nonmitochondrial); converts acetyl-CoA and oxaloacetate to citrate plus CoA
Eno1	Enolase 1 (2-phosphoglycerate dehydratase); converts 2-phospho-D-glycerate to phosphoenolpyruvate in glycolysis
Eno2	Enolase 2 (2-phosphoglycerate dehydratase); converts 2-phospho-D-glycerate to phosphoenolpyruvate in glycolysis
Fba1	Fructose bisphosphate aldolase II; sixth step in glycolysis
Hxk1	Hexokinase I; converts hexoses to hexose phosphates in glycolysis; repressed by glucose
Hxk2	Hexokinase II; converts hexoses to hexose phosphates in glycolysis and plays a regulatory role in glucose repression
Icl1	Isocitrate lyase, peroxisomal; carries out part of the glyoxylate cycle; required for gluconeogenesis
Pdb1	Pyruvate dehydrogenase complex, E1 beta subunit
Pdc1	Pyruvate decarboxylase isozyme 1
Pfk1	Phosphofructokinase alpha subunit, part of a complex with Pfk2p which carries out a key regulatory step in glycolysis
Pgi1	Glucose-6-phosphate isomerase, converts glucose-6-phosphate to fructose-6-phosphate
Pyc1	Pyruvate carboxylase 1; converts pyruvate to oxaloacetate for gluconeogenesis
Tal1	Transaldolase; component of nonoxidative part of pentose phosphate pathway
Tdh2	Glyceraldehyde-3-phosphate dehydrogenase 2; converts D-glyceraldehyde 3-phosphate to 1,3-dephosphoglycerate
Tdh3	Glyceraldehyde-3-phosphate dehydrogenase 3; converts D-glyceraldehyde 3-phosphate to 1,3-dephosphoglycerate
Tpi1	Triosephosphate isomerase; interconverts glyceraldehyde-3-phosphate and dihydroxyacetone phosphate
Efb1	Translation elongation factor EF-1β; GDP/GTP exchange factor for Tef1p/Tef2p
Eft1	Translation elongation factor EF-2; contains diphthamide which is not essential for activity; identical to Eft2p
Eft2	Translation elongation factor EF-2; contains diphthamide which is not essential for activity; identical to Eft1p
Prt1	Translation initiation factor eIF3 beta subunit (p90); has an RNA recognition domain
Rpa0 (RPPO)	Acidic ribosomal protein A0
Tif1	Translation initiation factor 4A (eIF4A) of the DEAD box family
Tif2	Translation initiation factor 4A (eIF4A) of the DEAD box family
Yef3	Translation elongation factor EF-3A; member of ATP-binding cassette superfamily
Hsc82	Chaperonin homologous to *E. coli* HtpG and mammalian HSP90
Hsp60	Mitochondrial chaperonin that cooperates with Hsp10p; homolog of *E. coli* GroEL
Hsp82	Heat-inducible chaperonin homologous to *E. coli* HtpG and mammalian HSP90
Hsp104	Heat shock protein required for induced thermotolerance and for resolubilizing aggregates of denatured proteins; important for $[psi^-]$-to-$[PSI^+]$ prion conversion
Kar2	Heat shock protein of the endoplasmic reticulum lumen required for protein translocation across the endoplasmic reticulum membrane and for nuclear fusion; member of the HSP70 family
Ssa1	Cytoplasmic chaperone; heat shock protein of the HSP70 family
Ssa2	Cytoplasmic chaperone; member of the HSP70 family
Ssb1	Heat shock protein of HSP70 family involved in the translational apparatus
Ssb2	Heat shock protein of HSP70 family, cytoplasmic
Ssc1	Mitochondrial protein that acts as an import motor with Tim44p and plays a chaperonin role in receiving and folding of protein chains during import; heat shock protein of HSP70 family
Sse1	Heat shock protein of the HSP70 family; multicopy suppressor of mutants with hyperactivated Ras/cyclic AMP pathway
Sti1	Stress-induced protein required for optimal growth at high and low temperature; has tetratricopeptide repeats
Ade1	Phosphoribosylamidoimidazole-succinocarboxamide synthase: catalyzes the seventh step in de novo purine biosynthesis pathway
Ade3	C_1 tetrahydrofolate synthase (trifunctional enzyme), cytoplasmic
Ade5,7	Phosphoribosylamine-glycine ligase plus phosphoribosylformylglycinamidine cyclo-ligase; bifunctional protein
Arg4	Argininosuccinate lyase; catalyzes the final step in arginine biosynthesis
Gdh1	Glutamate dehydrogenase ($NADP^+$); combines ammonia and α-ketoglutarate to form glutamate
Gln1	Glutamine synthetase; combines ammonia to glutamate in ATP-driven reaction
His4	Phosphoribosyl-AMP cyclohydrolase/phosphoribosyl-ATP pyrophosphohydrolase/histidinol dehydrogenase; 2nd, 3rd, and 10th steps of his biosynthesis pathway
Ilv5	Ketol-acid reductoisomerase (acetohydroxy, acid reductoisomerase) (alpha-keto-β-hydroxylacyl) reductoisomerase); second step in Val and Ilv biosynthesis pathway
Lys9	Saccharopine dehydrogenase ($NADP^+$, L-glutamate forming) (saccharopine reductase), seventh step in lysine biosynthesis pathway
Met6	Homocysteine methyltransferase; (5-methyltetrahydropteroyl triglutamate-homocysteine methyltransferase), methionine synthase, cobalamin independent
Pro2	γ-Glutamyl phosphate reductase (phosphoglutamate dehydrogenase), proline biosynthetic enzyme
Ser1	Phosphoserine transaminase; involved in synthesis of serine from 3-phosphoglycerate
Trp5	Tryptophan synthase, last (5th) step in tryptophan biosynthesis pathway
Act1	Actin; involved in cell polarization, endocytosis, and other cytoskeletal functions
Adk1	Adenylate kinase (GTP:AMP phosphotransferase), cytoplasmic
Ald6	Cytosolic acetaldehyde dehydrogenase
Atp2	Beta subunit of F1-ATP synthase; 3 copies are found in each F1 oligomer
Bmh1	Homolog of mammalian 14-3-3 protein; has strong similarity to Bmh2p
Bmh2	Homolog of mammalian 14-3-3 protein; has strong similarity to Bmh1p
Cdc48	Protein of the AAA family of ATPases; required for cell division and homotypic membrane fusion
Cdc60	Leucyl-tRNA synthetase, cytoplasmic
Erg20	Farnesyl pyrophosphate synthetase; may be rate-limiting step in sterol biosynthesis pathway
Gpp1 (Rhr2)	DL-Glycerol phosphate phosphatase
Gsp1	Ran, a GTP-binding protein of the Ras superfamily involved in trafficking through nuclear pores
Ipp1	Inorganic pyrophosphatase, cytoplasmic
Lcb1	Component of serine C-palmitoyltransferase; first step in biosynthesis of long-chain base component of sphingolipids
Mol1 (Thi4)	Thiamine-repressed protein essential for growth in the absence of thiamine
Pab1	Poly(A)-binding protein of cytoplasm and nucleus; part of the 3′-end RNA-processing complex (cleavage factor I); has 4 RNA recognition domains
Psa1	Mannose-1-phosphate guanyltransferase; GDP-mannose pyrophosphorylase
Rnr4	Ribonucleotide reductase small subunit
Sam1	*S*-Adenosylmethionine synthetase 1
Sam2	*S*-Adenosylmethionine synthetase 2
Sod1	Copper-zinc superoxide dismutase
Uba1	Ubiquitin-activating (E1) enzyme
YKL056	Resembles translationally controlled tumor protein of animal cells and higher plants
YLR109 (Ahp1)	Alkyl hydroperoxide reductase
YMR116 (Asc1)	Abundant protein with effects on translational efficiency and cell size, has two WD (WD-40) repeats

[a] Accepted name from the *Saccharomyces* genome database and YPD. Names in parentheses represent recent changes.
[b] Courtesy of Proteome, Inc., reprinted with permission.

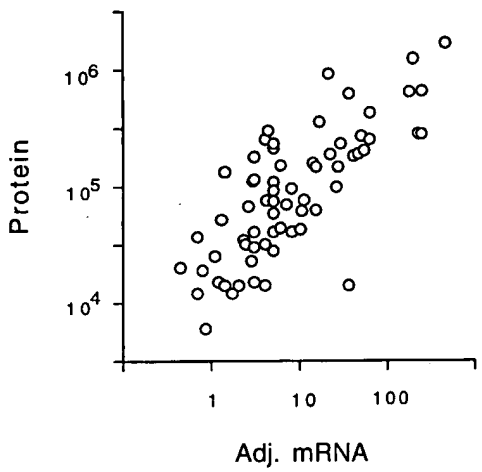


FIG. 2. Correlation of protein abundance with adjusted mRNA abundance. The number of molecules per cell of each protein is plotted against the number of molecules per cell of the cognate mRNA, with an r_p of 0.76. Note the logarithmic axes. Data for mRNA were taken from references 27 and 30 and combined as described in Materials and Methods.

correlation, confirming a strong correlation between CAI and protein abundance (Fig. 3). The relationship between CAI and protein abundance is log linear from about 1,000,000 to about 10,000 molecules per cell. We have no data for rarer proteins.

It is not clear whether CAI reflects maximum or average levels of protein expression. The proteins used for the CAI-protein correlation included some proteins which were not expressed at maximum levels under the condition of the experiment (Hsc82, Hsp104, Ssa1, Ade1, Arg4, His4, and others). When these proteins were removed from consideration and the correlation between CAI and the remaining (presumably constitutive) proteins was recalculated, the r_s was essentially unchanged (not shown).

The equation describing the graph in Fig. 3 is log (protein molecules/cell) = (2.3 × CAI) + 3.7. Thus, under certain conditions (a CAI of 0.3 or greater; a constitutively expressed gene), a very rough estimate of protein abundance can be made by raising 10 to the power of [(2.3 × CAI) + 3.7].

The distribution of CAI over the genome (Fig. 4) consists of a lower, bell-shaped distribution, possibly indicating a region where there is no selection for codon bias, and an upper, flat distribution, starting at a CAI of about 0.3, possibly indicating a region where there is selection for codon bias. Almost all of the proteins whose abundance we have measured are in the upper, flat portion of the distribution. In the lower, bell-shaped region, we do not know whether there is a correlation between CAI and protein abundance.

Changes in protein abundance in glucose and ethanol. A comparison of cells grown in glucose (Fig. 1A) with cells grown in ethanol (Fig. 1B) is shown in Table 1. As is well known, some proteins are induced tremendously during growth on ethanol. Two striking examples are the peroxisomal enzymes Icl1 (isocitrate lyase) and Cit2 (citrate synthase), which are induced in ethanol by more than 100- and 12-fold, respectively (Fig. 1; Table 1). These enzymes are key components of the glyoxylate shunt, which diverts some acetyl coenzyme A (acetyl-CoA) from the tricarboxylic acid cycle to gluconeogenesis. *S. cerevisiae* requires large amounts of carbohydrate for its cell wall; in ethanol medium, this carbohydrate comes from gluconeogenesis, which depends on the glyoxylate shunt and on the glycolytic pathway running in reverse. The need for gluconeogenesis also explains why glycolytic enzymes are abundant even in ethanol medium. Thus, 2D gel analysis shows the prominence of the glycolytic and glyoxylate shunt enzymes in cells grown on ethanol, emphasizing that gluconeogenesis, presumably largely for production of the cell wall, is a major metabolic activity under these conditions.

During gluconeogenesis, substrate-product relationships are reversed for the glycolytic enzymes. One might expect that not all glycolytic enzymes would be well adapted to the reverse reaction. Indeed, 2D gels show that in ethanol, Adh2 (alcohol dehydrogenase 2) is strongly induced (16), while its isozyme Adh1 is not greatly affected. Adh1 and Adh2 each interconvert acetaldehyde and ethanol. Adh1 has a relatively high K_m for ethanol (17 mM), while Adh2 has a lower K_m (0.8 mM) (5). Thus, it is thought that Adh1 is specialized for glycolysis (acetaldehyde to ethanol), while Adh2 is specialized for respiration (ethanol to acetaldehyde) (5, 29). Similarly, Eno1 (enolase 1) is induced in ethanol, while its isozyme Eno2 (enolase 2) decreases in abundance (Table 1) (4, 19). Eno1 is inhibited by 2-phosphoglycerate (the glycolytic substrate), while Eno2 is inhibited by phosphoenolpyruvate (the gluconeogenic substrate) (4). Perhaps Eno1 has a lower K_m for phosphoenolpyruvate than does Eno2, though to our knowledge this has not been tested. Thus, the 2D gels distinguish isozymes specialized for growth on glucose (Adh1 and Eno2) from isozymes specialized for ethanol (Adh2 and Eno1).

Many heat shock proteins (e.g., Hsp60, Hsp82, Hsp104, and Kar2) were about twofold more abundant in ethanol medium than in glucose medium. This is consistent with the increased heat resistance of cells grown in ethanol (3).

Enzymes involved in protein synthesis (Eft1, Rpa0, and Tif1) were about twice as abundant in glucose medium as in ethanol medium. This may reflect the higher growth rate of the cells in glucose.

Phosphorylation of proteins. To examine protein phosphorylation, we labeled cells with ^{32}P and ran 2D gels to examine phosphoproteins. About 300 distinct spots, probably representing 150 to 200 proteins, could be seen on pH 4–8 gels (Fig. 5B). We then aligned autoradiograms of three gels, each with a different kind of labeled protein (^{32}P only [Fig. 5B], ^{32}P plus ^{35}S [Fig. 5A], and ^{35}S only [not shown, but see Fig. 1 for example]). In this way, we made provisional identification of

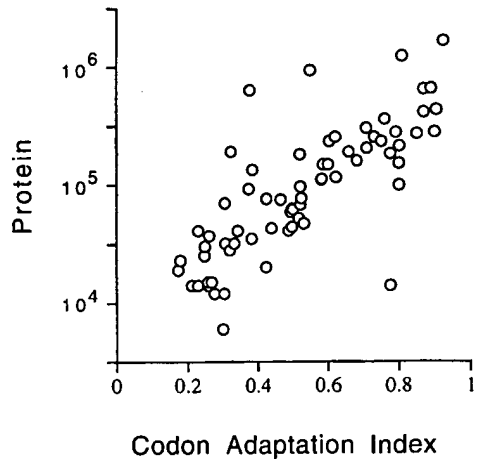


FIG. 3. Correlation of protein abundance with CAI. The number of molecules per cell of each protein is plotted against the CAI for that protein. Note the logarithmic scale on the protein axis. Data for the CAI are from the YPD database (13).

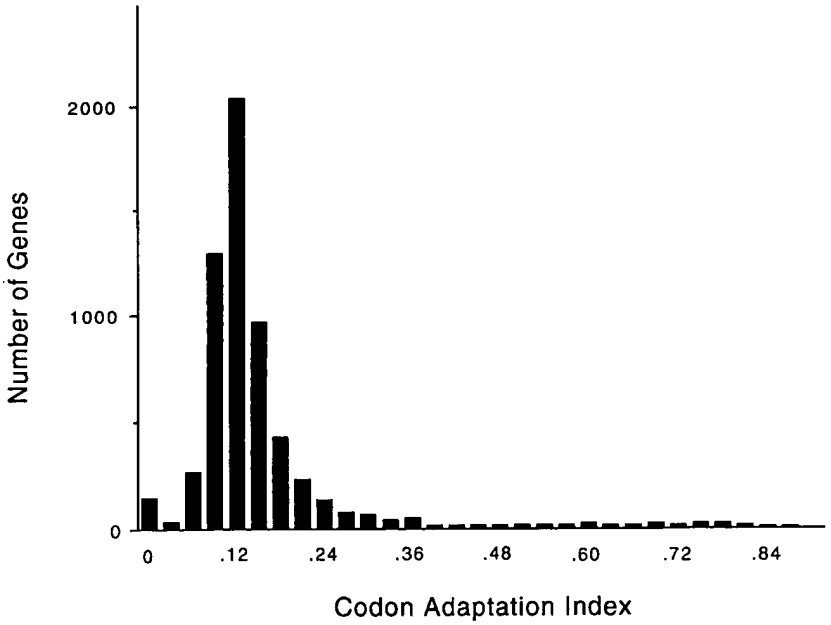


FIG. 4. Distribution of CAI over the whole genome, shown in intervals of 0.030 (i.e., there are 150 genes with a CAI between 0.000 and 0.030, inclusive; 31 genes with a CAI between 0.031 and 0.060; 269 genes with a CAI between 0.061 and 0.090; 1,296 genes with a CAI between 0.091 and 0.120; etc.). The distribution peaks with 2,028 genes with a CAI between 0.121 and 0.150.

some of the ^{32}P-labeled spots as particular ^{35}S-labeled spots. All such identifications are somewhat uncertain, since precise alignments are difficult, and of course multiple spots may exactly comigrate. Nevertheless, we believe that most of the provisional identifications are probably correct. Among the major ^{32}P-labeled proteins are the hexokinases Hxk1 and Hxk2, the acidic ribosome-associated protein Rpa0, the translation factors Yef3 and Efb1, and probably Hsp70 heat shock proteins of the Ssa and Ssb families. Rpa0 and Efb1 are quantitatively monophosphorylated.

Many yeast proteins resolve into multiple spots on these 2D gels (7). Yef3 has five or more spots, at least four of which comigrate with ^{32}P. Tpi1 has a major spot showing no ^{32}P labeling and a minor, more acidic spot which overlaps with some ^{32}P label. Tif1 has at least seven spots (7); two of these overlap with some ^{32}P label, but five do not (Fig. 5). Eft1 has at least three spots (7), and none of these overlap with ^{32}P, although there are three nearby, unidentified ^{32}P-labeled spots (a, c, and d in Fig. 5). Spots that seem to be extra forms of Met6, Pdc1, Eno2, and Fba1 can be seen in Fig. 6A, but there is little ^{32}P at these positions in Fig. 5. Thus, phosphorylation explains some but not all of the different protein isoforms seen.

The cell cycle is regulated in part by phosphorylation. We compared ^{32}P-labeled proteins from cells synchronized in G_1 with α-factor, in cells synchronized in G_1 by depletion of G_1 cyclins, and in cells synchronized in M phase with nocodazole. Only very minor differences were seen, and these were difficult to reproduce. The cell cycle proteins regulated by phosphorylation may not be abundant enough for this technique to be applied easily.

Centrifugal fractionation. We fractionated ^{35}S-labeled extracts by centrifugation (Materials and Methods). Figure 6A shows the proteins in the supernatant of a high-speed (100,000 × *g*, 30 min) centrifugation, while Fig. 6B shows the proteins in the pellet of a low-speed (16,000 × *g*, 10 min) centrifugation. Many proteins are tremendously enriched in one fraction or the other, while others are present in both. Most glycolytic enzymes (e.g., Tdh2, Tdh3, Eno2, Pdc1, Adh1, and Fba1) are enriched in the supernatant fraction. The only exception is Pfk1 (not indicated), which is found in both pellet and supernatant fractions. Many proteins involved in protein synthesis (Eft1, Yef3, Prt1, Tif1, and Rpa0) are in the pellet, possibly because of the association of ribosomes with the endoplasmic reticulum. However, Efb1 is in the supernatant, as is a substantial portion of the Eft1. Perhaps surprisingly, several mitochondrial proteins (Atp2 [not shown] and Ilv5) are largely in the supernatant. Perhaps glass bead breakage of cells releases mitochondrial proteins. The nuclear protein Gsp1 is in the pellet fraction. The enrichment produced by centrifugation makes it possible to see minor spots which are otherwise poorly resolved from surrounding proteins. Figure 6B shows that the previously identified Tif1 spot is surrounded by as many as six other spots that cofractionate. We observed six identical or very similar additional spots when we overexpressed Tif1 from a high-copy-number plasmid (not shown). Signal overlaps only one or two of these spots in ^{32}P-labeling experiments (Fig. 5), and so the different forms are not mainly due to different phosphorylation states.

DISCUSSION

Our experience with developing a 2D gel protein database for *S. cerevisiae* is summarized here. With current technology, we can see the most abundant 1,200 proteins, which is about one-third to one-quarter of the proteins expressed. The remaining proteins will be difficult to see and study with the methods that we have used, not because of a lack of sensitivity but because weak spots are covered by nearby strong spots.

Of the 1,200 proteins seen, we have identified 148, with a bias toward the most abundant proteins. Steady application of the methods already used would allow identification of most of the remaining proteins. Gene overexpression will be particularly useful, since it is not affected by the lower abundance of the remaining visible proteins.

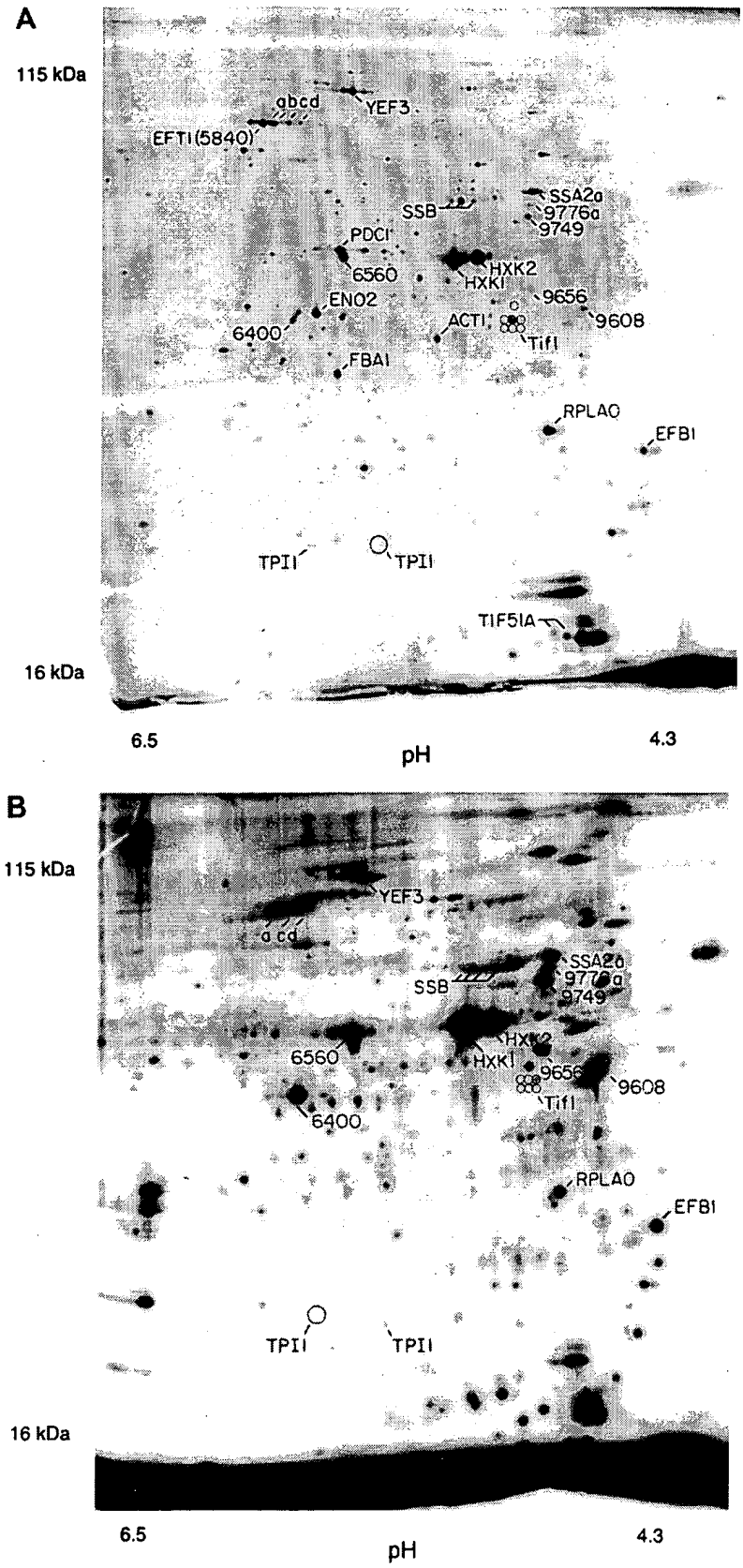


FIG. 5. Phosphorylated proteins. (A) Mixture of ^{32}P-labeled proteins and ^{35}S-labeled proteins. Two separate labeling reactions were done, one with ^{32}P and one with ^{35}S, and extracts were mixed and run on a 2D gel. Spots marked with numbers rather than gene names represent spots noted on ^{35}S gels but unidentified. Spots labeling with ^{32}P were identified by (i) increased labeling compared to the ^{35}S-only gel (not shown); (ii) the characteristic fuzziness of a ^{32}P-labeled spot; and (iii) the decay of signal intensity seen on exposures made 4 weeks later (not shown). A minor form of Tpi1 and at least six minor forms of Tif1 have been noted in overexpression experiments (see also Fig. 6B); positions of the minor forms are indicated by circles. (B) ^{32}P-only labeling. The major form of Tpi1, which is not labeled with ^{32}P, is indicated by a large circle; positions of seven forms of Tif1 are indicated by smaller circles.

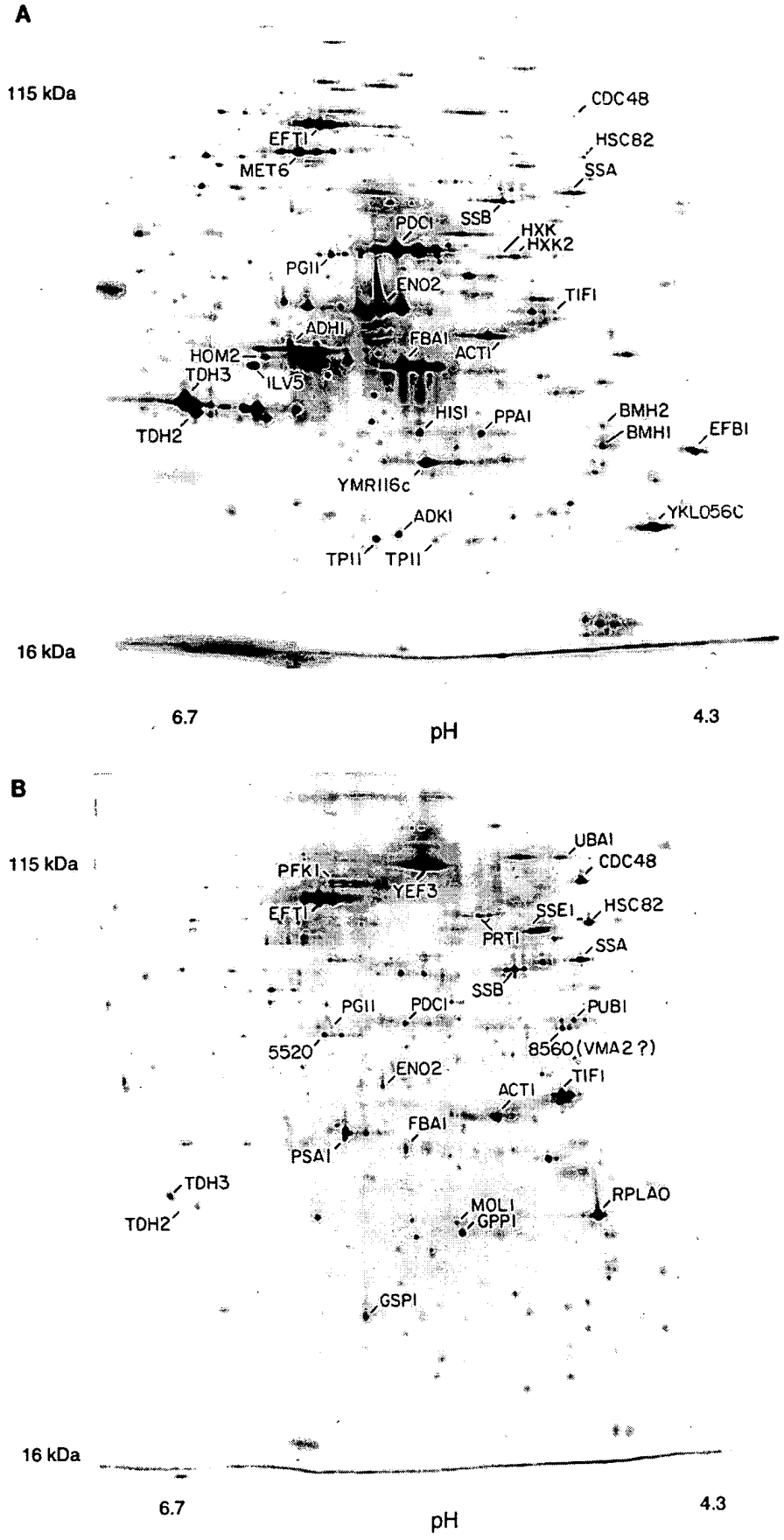


FIG. 6. Fractionation by centrifugation. (A) Proteins in the supernatant of a 100,000 × *g*, 30-min spin; proteins in the pellet of a 16,000 × *g*, 10-min spin. Supernatant fractions examined in multiple experiments done over a wide range of *g* forces looked similar to each other, as did the pellet fractions.

2D gels of the kind that we have used are not suitable for visualization of rare proteins. However it will be possible to study on a global basis metabolic processes involving relatively abundant proteins, such as protein synthesis, glycolysis, gluconeogenesis, amino acid synthesis, cell wall synthesis, nucleotide synthesis, lipid metabolism, and the heat shock response.

Gygi et al. (10) have recently completed a study similar to ours. Despite generating broadly similar data, Gygi et al. reached markedly different conclusions. We believe that both mRNA abundance and codon bias are useful predictors of protein abundance. However, Gygi et al. feel that mRNA abundance is a poor predictor of protein abundance and that "codon bias is not a predictor of either protein or mRNA levels" (10). These different conclusions are partly a matter of viewpoint. Gygi et al. focus on the fact that the correlations of mRNA and codon bias with protein abundance are far from perfect, while we focus on the fact that, considering the wide range of mRNA and protein abundance and the undoubted presence of other mechanisms affecting protein abundance, the correlations are quite good.

However, the different conclusions are also partly due to different methods of statistical analysis and to real differences in data. With respect to statistics, Gygi et al. used the Pearson product-moment correlation coefficient (r_p) to measure the covariance of mRNA and protein abundance. Depending on the subset of data included, their r_p values ranged from 0.1 to 0.94. Because of the low r_p values with some subsets of the data, Gygi et al. concluded that the correlation of mRNA to protein was poor. However, the r_p correlation is a parametric statistic and so requires variates following a bivariate normal distribution; that is, it would be valid only if both mRNA and protein abundances were normally distributed. In fact, both distributions are very far from normal (data not shown), and so a calculation of r_p is inappropriate. There was no statistical backing for the assertion that codon bias fails to predict protein abundance.

We have taken two statistical approaches. First, we have used the Spearman rank correlation coefficient (r_s). Since this statistic is nonparametric, there is no requirement for the data to be normally distributed. Using the r_s, we find that mRNA abundance is well correlated with protein abundance (r_s = 0.74), and the CAI is also well correlated with protein abundance (r_s = 0.80) (and also with mRNA abundance [data not shown]). For the data of Gygi et al. (10), we obtained similar results, though with their data the correlation is not as good; r_s = 0.59 for the mRNA-to-protein correlation, and r_s = 0.59 for the codon bias-to-protein correlation.

In a second approach, we transformed the mRNA and protein data to forms where they were normally distributed, to allow calculation of an r_p (Materials and Methods). Two transformations, Box-Cox and logarithmic, were used; both gave good correlations with our data [e.g., r_p = 0.76 for log(adjusted RNA) to log(protein)]. We were not able to transform the data of Gygi et al. to a normal distribution.

Finally, there are also some differences in data between the two studies. These may be partly due to the different measurement techniques used: Gygi et al. measured protein abundance by cutting spots out of gels and measuring the radioactivity in each spot by scintillation counting, whereas we used phosphorimaging of intact gels coupled to image analysis. We compared our data to theirs for the proteins common between the studies (but excluding proteins whose mRNAs are known to differ between rich and minimal media, and excluding Tif1, which was anomalous in differing by 100-fold between the two data sets). The r_s between the two protein data sets was 0.88 ($P < 0.0001$). Although this is a strong correlation, the fact that it is less than 1.0 suggests that there may have been errors in measuring protein abundance in one or both studies. After normalizing the two data sets to assume the same amount of protein per cell, we found a systematic tendency for the protein abundance data of Gygi et al. to be slightly higher than ours for the highest-abundance proteins and also for the lowest-abundance proteins but slightly lower than ours for the middle-abundance proteins. These systematic differences suggest some systematic errors in protein measurement. Although we do not know what the errors are, we suggest the following as a reasonable speculation. For the highest-abundance proteins, we may have underestimated the amount of protein because of a slightly nonlinear response of the phosphorimager screens. For the lowest-abundance proteins, Gygi et al. may have overestimated the amount of protein because of difficulties in accurately cutting very small spots out of the gel and because of difficulties in background subtraction for these small, weak spots. The difference in the middle abundance proteins may be a consequence of normalization, given the two errors above.

The low-abundance proteins in the data set of Gygi et al. have a poor correlation with mRNA abundance. We calculate that the r_s is 0.74 for the top 54 proteins of Gygi et al. but only 0.22 for the bottom 53 proteins, a statistically significant difference. However, with our data set, the r_s is 0.62 for the top 33 proteins and 0.56 (not significantly different) for the bottom 33 proteins (which are comparable in abundance to the bottom 53 proteins of Gygi et al.). Thus, our data set maintains a good correlation between mRNA and protein abundance even at low protein abundance. This is consistent with our speculation that protein quantification by phosphorimaging and image analysis may be more accurate for small, weak spots than is cutting out spots followed by scintillation counting. Our relatively good correlations even for nonabundant proteins may also reflect the fact that we used both SAGE data and RNA hybridization data, which is most helpful for the least abundant mRNAs. In summary, we feel that the poor correlation of protein to mRNA for the nonabundant proteins of Gygi et al. may reflect difficulty in accurately measuring these nonabundant proteins and mRNAs, rather than indicating a truly poor correlation in vivo. It is not surprising that observed correlations would be poorer with less-abundant proteins and mRNAs, simply because the accuracy of measurement would be worse.

How well can mRNA abundance predict protein abundance? With r_p = 0.76 for logarithmically transformed mRNA and protein data, the coefficient of determination, $(r_p)^2$, is 0.58. This means that more than half (in log space) of the variation in protein abundance is explained by variation in mRNA abundance. When converted back to arithmetic values, protein abundances vary over about 200-fold (Table 1), and $(r_p)^2$ = 0.58 for the log data means that of this 200-fold variation, about 20-fold is explained by variation in the abundance of mRNA and about 10-fold is unexplained (but could be due partly to measurement errors). For proteins much less abundant than those considered here, we imagine the *in vivo* correlation between mRNA and protein abundance will be worse, and other regulatory mechanisms such as protein turnover will be more important.

Some important conclusions can be drawn from this sampling of the proteome. First, there is an enormous range of protein abundance, from nearly 2,000,000 molecules per cell for some glycolytic enzymes to about 100 per cell for some cell cycle proteins (26a). Second, about half of all cellular protein is found in fewer than 100 different gene products, which are mostly involved in carbohydrate metabolism or protein synthe-

sis. Third, the correlation between protein abundance and CAI is log linear as far as we can see, which is from about 10,000 protein molecules per cell to about 1,000,000. This is somewhat surprising, because it implies that selective forces for codon bias are significant even at moderate expression levels. It also means that codon bias is a useful predictor of protein abundance even for moderately low bias proteins. Fourth, there is a good correlation between protein abundance and mRNA abundance for the proteins that we have studied. This validates the use of mRNA abundance as a rough predictor of protein abundance, at least for relatively abundant proteins. Fifth, for these abundant proteins, there are about 4,000 molecules of protein for each molecule of mRNA. This last conclusion raises questions as to how the levels of nonabundant proteins are regulated and suggests that protein instability, regulated translation, suboptimal rates of translation, and other mechanisms in addition to transcriptional control may be very important for these proteins.

ACKNOWLEDGMENTS

We thank Neena Sareen and Nick Bizios (CSHL 2D gel laboratory) for production of 2D gels, Tom Volpe for help with some experiments, Corine Driessens for help with calculations and statistics, and Herman Wijnen and Nick Edgington for comments on the manuscript. We especially thank Tim Tully for in-depth statistical analysis and for insightful discussions on statistical interpretations.

This work was supported by grant P41-RR02188 from the NIH Biomedical Research Technology Program, Division of Research Resources, to J.I.G., by Small Business Innovation Research grant R44 GM54110 to Proteome, Inc., by grant DAMD17-94-J4050 from the Army Breast Cancer Program to B.F., and by NIH grant RO1 GM45410 to B.F.

REFERENCES

1. **Baroni, M. D., E. Martegani, P. Monti, and L. Alberghina.** 1989. Cell size modulation by *CDC25* and *RAS2* genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 9:2715–2723.
2. **Boucherie, H., F. Sagliocco, R. Joubert, I. Maillet, J. Labarre, and M. Perrot.** 1996. Two-dimensional gel protein database of Saccharomyces cerevisiae. Electrophoresis 17:1683–1699.
3. **Elliott, B., and B. Futcher.** 1993. Stress resistance of yeast cells is largely independent of cell cycle phase. Yeast 9:33–42.
4. **Entian, K. D., B. Meurer, H. Kohler, K. H. Mann, and D. Mecke.** 1987. Studies on the regulation of enolases and compartmentation of cytosolic enzymes in Saccharomyces cerevisiae. Biochim. Biophys. Acta 923:214–221.
5. **Ganzhorn, A. J., D. W. Green, A. D. Hershey, R. M. Gould, and B. V. Plapp.** 1987. Kinetic characterization of yeast alcohol dehydrogenases. Amino acid residue 294 and substrate specificity. J. Biol. Chem. 262:3754–3761.
6. **Garrels, J. I.** 1989. The Quest system for quantitative analysis of two-dimensional gels. J. Biol. Chem. 264:5269–5282.
7. **Garrels, J. I., B. Futcher, R. Kobayashi, G. I. Latter, B. Schwender, T. Volpe, J. R. Warner, and C. S. McLaughlin.** 1994. Protein identifications for a Saccharomyces cerevisiae protein database. Electrophoresis 15:1466–1486.
8. **Garrels, J. I., C. S. McLaughlin, J. R. Warner, B. Futcher, G. I. Latter, R. Kobayashi, B. Schwender, T. Volpe, D. S. Anderson, R. Mesquita-Fuentes, and W. E. Payne.** 1997. Proteome studies of S. cerevisiae: identification and characterization of abundant proteins. Electrophoresis 18:1347–1360.
9. **Goffeau, A., B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, and S. G. Oliver.** 1996. Life with 6000 genes. Science 274:563–567.
10. **Gygi, S. P., Y. Rochon, B. R. Franza, and R. Aebersold.** 1999. Correlation between protein and mRNA abundance in yeast. Mol. Cell. Biol. 19:1720–1730.
11. **Hereford, L. M., and M. Rosbash.** 1977. Number and distribution of polyadenylated RNA sequences in yeast. Cell 10:453–462.
12. **Herrick, D., R. Parker, and A. Jacobson.** 1990. Identification and comparison of stable and unstable mRNAs in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 10:2269–2284.
13. **Hodges, P. E., A. H. McKee, B. P. Davis, W. E. Payne, and J. I. Garrels.** 1999. The Yeast Proteome Database (YPD): a model for the organization of genome-wide functional data. Nucleic Acids Res. 27:69–73.
14. **Ikemura, T.** 1985. Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. 2:13–34.
15. **Johnston, G. C., F. R. Pringle, and L. H. Hartwell.** 1977. Coordination of growth with cell division in the yeast S. cerevisiae. Exp. Cell Res. 105:79–98.
16. **Johnston, M., and M. Carlson.** 1992. Regulation of carbon and phosphate utilization, p. 193–281. *In* E. Jones, J. Pringle, and J. Broach (ed.), The molecular and cellular biology of the yeast Saccharomyces. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
17. **Kornblatt, M. J., and A. Klugerman.** 1989. Characterization of the enolase isozymes of rabbit brain: kinetic differences between mammalian and yeast enolases. Biochem. Cell. Biol. 67:103–107.

17a. **Latter, G., and B. Futcher.** Unpublished data.

18. **Mathews, B., N. Sonenberg, and J. W. B. Hershey.** 1996. Origins and targets of translational control, p. 1–29. *In* J. W. B. Hershey, M. B. Mathews, and N. Sonenberg (ed.), Translational control. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
19. **McAlister, L., and M. J. Holland.** 1982. Targeted deletion of a yeast enolase structural gene. Identification and isolation of yeast enolase isozymes. J. Biol. Chem. 257:7181–7188.
20. **Monardo, P. J., T. Boutell, J. I. Garrels, and G. I. Latter.** 1994. A distributed system for two-dimensional gel analysis. Comput. Appl. Biosci. 10:137–143.
21. **O'Farrell, P. H.** 1975. High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. 250:4007–4021.
22. **Patterson, S. D., and G. I. Latter.** 1993. Evaluation of storage phosphor imaging for quantitative analysis of 2-D gels using the Quest II system. BioTechniques 15:1076–1083.
23. **Sagliocco, F., J. C. Guillemot, C. Monribot, J. Capdevielle, M. Perrot, E. Ferran, P. Ferrara, and H. Boucherie.** 1996. Identification of proteins of the yeast protein map using genetically manipulated strains and peptide-mass fingerprinting. Yeast 12:1519–1533.
24. **Sharp, P. M., and W. H. Li.** 1987. The Codon Adaptation Index—a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. 15:281–1295.
25. **Shevchenko, A., O. N. Jensen, A. V. Podtelejnikov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortensen, A. Shevchenko, H. Boucherie, and M. Mann.** 1996. Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc. Natl. Acad. Sci. USA 93:14440–14445.
26. **Thomas, B. J., and R. Rothstein.** 1989. Elevated recombination rates in transcriptionally active DNA. Cell 56:619–630.

26a. **Tyers, M., and B. Futcher.** Unpublished data.

27. **Velculescu, V. E., L. Zhang, W. Zhou, J. Vogelstein, M. A. Basrai, D. E. Bassett, Jr., P. Hieter, B. Vogelstein, and K. W. Kinzler.** 1997. Characterization of the yeast transcriptome. Cell 88:243–251.
28. **Warner, J.** 1991. Labeling of RNA and phosphoproteins in S. cerevisiae. Methods Enzymol. 194:423–428.
29. **Wills, C.** 1976. Production of yeast alcohol dehydrogenase isoenzymes by selection. Nature 261:26–29.

29a. **Wodicka, L.** Personal communication.

29b. **Wodicka, L.** Unpublished data.

30. **Wodicka, L., H. Dong, M. Mittmann, M.-H. Ho, and D. J. Lockhart.** 1997. Genome-wide expression monitoring in Saccharomyces cerevisiae. Nat. Biotechnol. 15:1359–1367.